ARL

US Army Research Laboratory

Simulation Training in Health Care

by Frank A Drews and Jonathan Z Bakdash

NOTICES

Disclaimers

The findings in this report are not to be construed as an official Department of the Army position unless so designated by other authorized documents.

Citation of manufacturer's or trade names does not constitute an official endorsement or approval of the use thereof.

Destroy this report when it is no longer needed. Do not return it to the originator.

ARL

US Army Research Laboratory

Simulation Training in Health Care

by Frank A Drews
University of Utah, Salt Lake City, UT
George Wahlen VA Medical Center, Salt Lake City, UT

and

Jonathan Z Bakdash
Human Research and Engineering Directorate, ARL

REPORT DOCUMENTATION PAGE		*Form Approved* *OMB No. 0704-0188*

Public reporting burden for this collection of information is estimated to average 1 hour per response, including the time for reviewing instructions, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing the collection information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing the burden, to Department of Defense, Washington Headquarters Services, Directorate for Information Operations and Reports (0704-0188), 1215 Jefferson Davis Highway, Suite 1204, Arlington, VA 22202-4302. Respondents should be aware that notwithstanding any other provision of law, no person shall be subject to any penalty for failing to comply with a collection of information if it does not display a currently valid OMB control number.
PLEASE DO NOT RETURN YOUR FORM TO THE ABOVE ADDRESS.

1. REPORT DATE *(DD-MM-YYYY)*	2. REPORT TYPE	3. DATES COVERED (From - To)
June 2015	Reprint	1 October 2012–30 September 2013

4. TITLE AND SUBTITLE	5a. CONTRACT NUMBER
Simulation Training in Health Care	
	5b. GRANT NUMBER
	5c. PROGRAM ELEMENT NUMBER
6. AUTHOR(S)	5d. PROJECT NUMBER
Frank A Drews and Jonathan Z Bakdash	
	5e. TASK NUMBER
	5f. WORK UNIT NUMBER

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)	8. PERFORMING ORGANIZATION REPORT NUMBER
US Army Research Laboratory ATTN: RDRL-HRS-E Aberdeen Proving Ground, MD 21005-5425	ARL-RP-0523

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)	10. SPONSOR/MONITOR'S ACRONYM(S)
	11. SPONSOR/MONITOR'S REPORT NUMBER(S)

12. DISTRIBUTION/AVAILABILITY STATEMENT

Approved for public release; distribution is unlimited.

13. SUPPLEMENTARY NOTES

A reprint from Reviews of Human Factors and Ergonomics. 2013;8:191–234.

14. ABSTRACT

In this report, we discuss the application of human factors and ergonomics to developing effective simulation training in health care. Simulation provides a safe, effective method for training and assessing human performance. In aviation, simulation-based training and assessment has been widely used, significantly improving safety. This progress would have been impossible without the involvement of human factors and ergonomics. Although aviation and health care have similarities, there also are differences that complicate the widespread implementation of simulation in health care.

15. SUBJECT TERMS

simulation, training, performance, patient safety, health care

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
					Jonathan Z Bakdash
a. REPORT	b. ABSTRACT	c. THIS PAGE	UU	50	19b. TELEPHONE NUMBER (Include area code)
Unclassified	Unclassified	Unclassified			410-278-9050

CHAPTER 5

Simulation Training in Health Care

Frank A. Drews & Jonathan Z. Bakdash

In this chapter, we discuss the application of human factors and ergonomics to developing effective simulation training in health care. Simulation provides a safe, effective method for training and assessing human performance. In aviation, simulation-based training and assessment has been widely used, significantly improving safety. This progress would have been impossible without the involvement of human factors and ergonomics. Although aviation and health care have similarities, there also are differences that complicate the widespread implementation of simulation in health care.

Simulation is an important element for improving patient safety. There is growing evidence that simulation training increases adherence to best practices, improves clinical outcomes, and reduces the costs associated with care. For effective simulation training in health care, or any domain, the simulation must support psychological fidelity, replicating the major cognitive operations of the actual or real-world (clinical) tasks. To achieve psychological fidelity, simulation development must be based on a thorough analysis of the tasks and goals present in the clinical environment. The simulation tool is selected with physical realism being only one means to support the psychological fidelity of the actual clinical task. This approach maximizes the efficacy of simulation training and minimizes costs.

Additional critical challenges to widespread implementation of effective simulation training in health care include the belief that high physical realism is necessary, a lack of standardization in clinical procedures and equipment, the absence of a mechanism for sharing simulation scenarios, and the difficulties associated with measuring clinical outcomes for patients with the consequent issues of validity and reliability of variables used as inputs in a cost-benefit analysis (CBA) of simulation training.

With the advancement of technology and the increasing need to improve human performance in complex systems, new approaches have emerged to support operators while performing complex tasks. One of these approaches is simulation-based training. Human factors and ergonomics is an important contributor to the creation of simulation based training programs (Jones, Hennessey, & Deutsch, 1985).

Today, simulations are used for diverse purposes, including entertainment, education, training, system evaluation, and research. They are commonly used in domains

where humans are required to manipulate or control complex systems, including aviation, health care, power plants, refineries, and many others. Some simulators are created to be as lifelike as possible and provide an experience nearly identical to the actual system, whereas others reproduce only the key characteristics of the simulated system. Regardless of their design or complexity, simulators are used in situations when use of the actual system is impractical because of financial limitations, system availability, ethical reasons, or risk of life (Haluck & Krummel, 2000; Schlectre, Bessemer, & Kolosh, 1992; Taylor, Lintern, & Koonce, 1993).

With effective training fundamental to the acquisition of the skills and knowledge necessary to deliver safe, high-quality health care, simulation training can support these goals. Traditional training in many health care areas involves a mentor–apprentice approach, whereby the trainee follows an experienced trainer and receives training with actual patients (Haluck & Krummel, 2000; Haque & Srinivasan, 2006). This practice is reflective of the still common training approach of "see one, do one, teach one" (Vozenilek, Huff, Reznek, & Gordon, 2004; Ziv, Wolpe, Small, & Glick, 2006). Gawande (2007) describes the consequences and the challenges of this approach.

The goal of this chapter is to provide a review that will be helpful to researchers currently using or planning to use simulation. In addition, practitioners interested in using a theory-driven and empirically supported approach to improve the quality of their simulations and measures to assess performance will benefit. Finally, human factors engineers and ergonomists with an interest in health care can benefit from reading this chapter as it focuses on improvement of health care delivery, one of the greatest and most interesting challenges to our field.

In this chapter, we review the current literature on simulation training in health care. First, we provide a brief overview of the history of simulation, ranging from aviation to health care. Second, we discuss the characteristics of simulation and the types of simulation that are being used for training and performance evaluation. This topic leads to the discussion of the rationale for simulator training in health care. In the fourth section of this chapter, we will analyze what makes a simulation effective, analyzing the psychological operations of the task with regard to the goals that are intended to be accomplished. In the fifth section, we describe the issue of cost analysis in the context of simulation. Last, the chapter concludes with a summary of the challenges of health care simulation and provides an outlook into the future of simulation in health care.

A BRIEF HISTORY OF SIMULATION: FROM AVIATION TO HEALTH CARE

The Beginnings: Simulation in Aviation

With the advancement of aviation at the beginning of the last century, frequent accidents of novice pilots marked the need for a better and safer approach toward training. In his discussion of virtual reality surgical simulation, Satava (2006, p. 2) summarizes this motivation well: "Whenever something is too dangerous, expensive,

or distant in time, place, or imagination to physically experience, there have been attempts to simulate the experience."

Moroney and Moroney (1999) and Page (2000) provide good overviews of the history of aviation simulation. Using these two references, we provide a brief overview here. At the beginning of the 20th century, interest in aviation was great, but training to become a pilot was expensive and dangerous. For example, early training required students to go through a sequence of exercises. After initial passenger flights, the student would attempt taxiing a low-powered machine to learn control of the rudder on the ground. Next, the student would advance to a plane that would allow small hops involving elevator control, which then would lead to actual flight.

In response to the high cost of flight training, alternate means for instruction emerged. Several engineers began the development of ground-based aviation trainers, which were essentially aircraft mounted to a universal joint that could be placed in a strong prevailing wind.

As a result of the high demand for pilots during World War I, the discipline of aviation psychology emerged with the goal of developing tools for more effective selection of future pilots. However, it was not until Edwin Link developed the Link Pilot Maker around 1928 that ground-based aviation simulators became widely used.

Link's Pilot Maker used air pressure to move actuators in response to control movements from the pilot, therefore removing the dependence on strong winds and manual movement of the previous training devices. When in 1934 the U.S. Army was searching for ways to teach instrument-only flight to pilots, it identified the Pilot Maker as a system that fulfilled its needs. Within a few years, Link Trainers were being sold to air forces across Europe and in Japan and became an invaluable tool to teach pilots to fly an airplane on instruments only (Smith & Smith, 1989).

The advancement of simulators took a leap due to the application of electrical and electronic methods. Because of the success of the training of aviation pilots with the Pilot Maker, other simulators were developed to train additional crew members. During the years that followed, simulators advanced in complexity, increased in size, and rose in cost. Following World War II, computers began to replace the levers, motors, and air bellows that previously controlled simulators. The simulator cockpits became near perfect replications of actual aircraft cockpits that included moving views of the ground. With advances in the technology behind the simulators, the scope of the training application increased as well. Simulators are now used to train myriad aspects of flight, including: basic system use (Bell & Waag, 1998; Stark, 1989), instrument use (Ortiz, 1994), decision making (Connolly, Blackwell, & Lester, 1989), and emergency procedures (Bell & Waag, 1998). The effective use of simulators in the field of aviation provided a foundation upon which other domains created their own simulation programs (Seymour et al., 2002).

Simulation in Health Care

Health care simulation can be dated back to the development of the part-task simulator Resusci-Anne (for training parts of a procedure) by the Laerdal Company in the 1950s. In response to a paper on the effectiveness of mouth-to-mouth ventilation

(Safar & McMahon, 1958), Resusci-Anne was developed to support skill acquisition for this procedure. With Resusci-Anne, Laerdal created a full-size training mannequin that allowed training on clinical scenarios, like obstructed airways. A later version of this simulator included a spring mechanism in the chest allowing simulation of chest compressions as well.

In the early 1960s, standardized or simulated patients (trained actors) were introduced to help train communication (e.g., taking a patient history) and clinical skills (e.g., physical examination) (Rosen, 2008), having the advantage of providing consistent conditions for assessing communication and clinical skills in specific situations (Wallace, 1997).

The next step in health care simulation was the development of a computer-controlled modern simulator (Abrahamson, Denson, & Wolf, 1969) in the late 1960s. Sim One was controlled by a hybrid digital and analog computer and had many features of today's high-physical-fidelity simulators: Sim One could breathe, had a heartbeat, displayed blood pressure, was able to blink, and could open and close its mouth. In addition, it was able to respond to four intravenously administered drugs and two gases. Sim One was utilized in studies of proficiency among anesthesiology residents. The evaluation of the effectiveness of this simulator indicated that simulator-trained residents reached professional levels of performance in fewer elapsed days and in a smaller number of trials in the operating room than those who did not participate in simulator training. One obstacle in the wide use of Sim One was its purchase and maintenance costs.

Approximately at the same time, the cardiopulmonary patient simulator Harvey was introduced at the University of Miami (http://www.gcrme.med.miami.edu/about_history_overview.php). Today's version of Harvey simulates cardiac and lung disease, including blood pressure, breathing, pulses, heart sounds, and murmurs.

A significant step toward the acceptance of medical simulation can be attributed to the work of Gaba at Stanford and of Good and Grevenstein at the University of Florida in the 1980s. Their work led to the creation of commercial medical simulators for use in medical training. Commercial availability made it easier to acquire and operate medical simulators, culminating in the establishment of simulation centers at medical schools nationwide. Today's popularity of simulation for medical education is illustrated by the fact that the number of annual simulation-related publications increased more than tenfold between the 1990s and the 2000s (Rosen, 2008). Excellent reviews of the history of medical simulation and a timeline of its development can be found in Grenvik and Schaefer (2004) and Rosen (2008), and a more focused discussion of the history of mannequin-based simulators can be found in Cooper and Taqueti (2004).

When comparing the development of simulation in aviation and in medicine, several similarities become apparent, leading to issues we will discuss in depth:

1. Simulation is a function of technology available; technological advances allow more "realistic" simulation. However, it is important to understand what exactly needs to be realistic in successful simulation training.

2. Technological advances lead to a broadening of the simulation training goals, exceeding the initial basic skills training. But it is not clear how to identify critical skills and how to develop effective simulation training.

3. The costs of simulation technology can be a limiting factor. Thus, it is important to clearly understand the economic costs and benefits of simulation use in training.

HOW TO DESCRIBE SIMULATION?

By modern standards, early aviation simulators, such as the Link Trainer, appear crude because they did not provide visual feedback nor did they use accurate physical modeling of aircraft aerodynamics. Nevertheless, such simulators effectively supported the psychological aspects of instrument-only flying. The underlying assumption for pursuing high physical realism in simulation is "… that by brute engineering force all physical and cognitive aspects of a task can be captured" (Caird, 1996, p. 127). There is a strong predilection toward highly realistic simulation (Scerbo & Dawson, 2007; Smallman & St. John, 2005). However, this preference is problematic because physical realism alone is neither necessary nor sufficient for effective training (Adams, 1979; Beaubien, 2004; Kozlowski & DeShon, 2004; Salas, Bowers, & Rhodenizer, 1998). Instead, it is vital that the simulation supports the psychological processes required for a specific task (Adams, 1979; Kozlowski & DeShon, 2004). For example, instrument-only flying does not necessarily require an exact model of physics and certainly not the visual simulation of weather conditions.

Typically, the term *fidelity* refers to physical fidelity. This definition is indicative of the common engineering approach to simulation. Simulator characteristics are multidimensional (related but partially dissociable) and on a continuum. We use *high* and *low* to refer to the admittedly somewhat subjective and ill-defined end points of physical and psychological fidelity. Other characteristics common to simulation include tractability, realism, and engagement. We use *realism* by itself to refer to the plausibility of events and action consequences in the simulation relative to the real world. These characteristics will now be discussed.

Simulator Characteristics

Fidelity. Stimulus fidelity refers to an exact match between the sensory output of the simulator and the actual system (Stoffregen, Bardy, Smart, & Pagulayan, 2003). Thus, a simulator with perfect fidelity would be indistinguishable from the actual system. However, complete physical realism and environment duplication in simulations may never be possible to achieve (Gibson, 1971; Hochberg, 1986; Stappers, Gaver, & Overbeeke, 2003; Stoffregen et al., 2003). Even though complete stimulus fidelity may not be achievable, the question is, is this is even necessary? For example, perceivers likely do not attend to all of the information provided to the same extent (Stoffregen et al., 2003), permitting some degradation of fidelity. One method to achieve high levels of simulator fidelity is

by use of model approximations and alternate dynamics (Stappers, 1997, 1998; Stappers & Waller, 1993).

An important consideration is the level of physical fidelity needed to be effective. A simulator used for training should have a level of fidelity required to successfully train the individuals using the simulator (Salas et al., 1998). Therefore, the development of simulation models requires a process of critical feature selection and quantitative description of relationships to accurately mimic the actual system. An effective simulation emulates important properties and maintains critical relationships while focusing on the purpose of the simulation.

In general, there is a preference for high visual realism by users and designers (Smallman & St. John, 2005). The conception of simulation development and design solely as an engineering problem and less as a knowledge transfer problem may explain the common inclination toward high physical realism. There are two potential disadvantages to high-fidelity simulations: cost and an absence of training benefits. High-fidelity simulation tools are generally much more expensive than low-fidelity simulations (Hopkins, 1974). In aviation, a high-physical-fidelity simulation does not necessarily result in more positive knowledge transfer than low-physical-fidelity simulation and sometimes even results in diminished transfer (Adams, 1979; Caird, 1996; Kozlowski & DeShon, 2004). The allure of high-fidelity simulation parallels a phenomenon in the graphical display of visual information: A visual display with more realistic graphics is preferred over one with less realistic graphics, even when the less realistic display leads to a better understanding of the information (Smallman & St. John, 2005). Smallman and St. John (2005) call this phenomenon "naive realism." We borrow the term *naive realism* to refer to high-physical-fidelity simulation that ignores psychological aspects of the training task.

Psychological and physical fidelity. Fidelity extends beyond the physical realm. Fidelity can also apply to subjective experience. Psychological and physical fidelity are partially related but also dissociable and encompass the two major dimensions of fidelity (see Rehmann, 1995 for subtypes of fidelity). Physical fidelity, also called engineering fidelity (Miller, 1957) or experiential fidelity (Stoffregen et al., 2003), is the degree to which the physical characteristics of the system are replicated within the simulation. Physical fidelity can be measured by the degree to which a simulation replicates the sensory characteristics of the actual environment. Psychological fidelity, also called action fidelity (Stoffregen et al., 2003), functional fidelity (Moroney & Moroney, 1999), or task fidelity (Rehmann, 1995), is the degree to which the simulation captures the psychological aspects of the real-world task or activity. Research has repeatedly shown that psychological fidelity is fundamental for transfer from the simulator to the real-world system (e.g., Kozlowski & DeShon, 2004).

Learning in a simulated environment does not necessarily require high physical fidelity. Effective training can often be accomplished with the use of low-fidelity simulators (Jentsch & Bowers, 1998; Koonce & Bramble, 1998; Patrick, 1992). This fact does not imply that high-fidelity simulation should always be avoided; instead, maintaining psychological fidelity to the task (based on the results of a task analysis, see Framework for

Effective Simulation Training section) should be used to determine the appropriate level of physical fidelity. To support psychological fidelity, some level of physical fidelity is necessary, but high physical realism alone is not sufficient. For example, if the task is suturing, a pig's foot has high physical fidelity for the properties of skin, a relevant aspect of the actual task (Barnes, Lang, & Whiteside, 1989). However, a pig's foot would not support psychological fidelity for providing the patient wound care instructions. An alternative to a pig's foot is a cadaver, because as an entire human body it has greater physical realism, although this physical realism may be irrelevant or even distracting to the task of suturing. These examples demonstrate that psychological fidelity must be supported by physical realism rather than general high fidelity.

Another example is simulating tissue handling in surgery. This task requires the knowledge, skills, and abilities to deliver the appropriate amount of force for cutting, pulling, and other actions (Al-Kadi et al., 2012). Training on virtual surgical simulators with haptic feedback leads to better instrument control and accuracy in clinical practice (Al-Kadi et al., 2012; Kim, Rattner, & Srinivasan, 2003). In addition to haptic feedback, adding computer animation to the surgical surfaces also improves skill acquisition because it is relevant to the perception of surface elasticity and other properties that matter for surgery (Basdogan et al., 2004). Although the fidelity of a simulator is generally reported in terms of the physical resemblance to the actual simulated system, its psychological fidelity is also critical to the successful application of simulation.

Tractability, realism, and engagement. A tractable simulation is one that can be readily used for its intended purpose. For example, researchers and professionals using simulators for training purposes are often caught among the overwhelming complexity of the field and the lack of complexity in laboratory contexts (Brehmer & Dörner, 1993). Tractable simulations allow the researcher to manage the level of complexity that is presented to the participant. When using simulation for training, tractability refers to the amount of training that users require before they can participate in the simulation, with training negatively correlated with tractability (Gray, 2002).

Realism is the extent to which actual experiences are encountered in the simulated system, including maintaining the functional relationships between interacting components of a system. For example, events that are likely to happen in the actual system will also occur in the simulated system. Realism differs from fidelity in that fidelity focuses on sensory stimulation, whereas realism focuses on plausible events and the consequences of actions that are executed.

Engagement relates to a participant's interest in and motivation toward maintaining high levels of performance throughout the scenario. It is the degree to which participants take the simulation seriously (Gray, 2002). Engagement can be improved by offering financial incentives for good performance, producing an interesting or enjoyable environment, and selecting participants with an interest in the subject matter (Ehret, Gray, & Kirschenbaum, 2000; Gray, 2002). Physical and psychological fidelity may help improve engagement, but the relationship between these characteristics is not well understood (see Dieckmann, Gaba, & Rall, 2007). We speculate that irrelevant physical realism (that does not support psychological fidelity) may

Figure 5.1. The 787 Simulator Flight Deck. Copyright © Boeing (left). Laerdal SimMan 3G high fidelity patient simulator located in a simulated operating room at the Anesthesiology Center for Patient Simulation at the University of Utah (right). Copyright 2013 by Noah Syroid.

increase engagement factors, such as motivation, yet negatively impact psychological fidelity by being potentially distracting.

There are deliberate distinctions in terms of tractability, realism, and engagement among different types of simulators. Research simulations typically require high levels of tractability and engagement, but they may not necessarily need to have high levels of physical realism. Training simulations may require high levels of relevant physical realism (in support of psychological fidelity) to allow appropriate learning of events, but they may not require high levels of tractability (e.g., initially, the clinical task is standardized) or engagement (e.g., learners may be intrinsically motivated) because skills can still be acquired without these.

Types of Simulators

Many simulators involve either complete or partial physical recreations of the simulated system, such as cockpits in flight simulators and patient simulators to medical procedures (see Figure 5.1). Some simulations deliberately exclude certain features of the simulated system to reduce complexity, whereas other simulations allow for component interaction that would be impossible in the real world (e.g., computer games). A simulation's application typically motivates its design. Because there are many diverse applications for simulators, each with unique requirements, there are a great variety of simulators. Five general classes of simulators can be identified: high fidelity, low fidelity, scaled worlds, microworlds, and virtual environments (VEs).

High-fidelity simulators. High-fidelity simulations have a close resemblance to the dynamics and appearance of the system that is being simulated. Driving simulators, such as the National Advanced Driving Simulator (NADS; www.nads-sc.uiowa.edu), have the identical controls, displays, and feel of an automobile since NADS uses a vehicle as part of the simulator. The graphical display extends to 360° of a photo-realistic

representation of the environment. In addition to realistically recreating the physical environment, many high-fidelity simulators attempt to mimic the complexity of the real world (Gray, 2002), such as replicating the behaviors of other drivers in traffic. High-fidelity simulations do not have to be complex, however. Realistic replications of simple systems can still have high fidelity without high complexity. Also, systems do not have to be replicated at full scale. Subsystems of complex systems can be created as stand-alone simulations (Gray, 2002). An example of a high-fidelity simulator currently used in medical simulation centers is the Human Patient Simulator (see Figure 5.1).

Low-fidelity simulators. Low-fidelity simulators do not have a close physical resemblance to the simulated system, but they may be rated high on psychological fidelity. Thus, it is possible that a simulator has low fidelity and still can maintain effectiveness. For example, numerous studies use a continuous-pursuit tracking task to simulate driving (e.g., Strayer & Johnston, 2001). An example of the use of a low-fidelity simulator in health care is the use of a pig's foot for suturing a wound.

Scaled worlds or part-task simulators. An alternate approach is to have a simulation that can help the learner to develop automaticity for individual parts of the task. Simulations of this type have been labeled as scaled worlds (Ehret et al., 2000) or part-task simulators (Perkins, 2007). Examples for this type of simulator can range from a central vascular catheter (CVC) insertion simulator to practice catheter insertion, to peg transfer tasks in a box simulator to simulate endoscopic/laparoscopic environment (Klein et al., 2008), to using an explanted pig's heart as an ex vivo part-task simulator for cardiac surgery (Fann et al., 2010). A part-task simulator provides an artificial environment that focuses on a subset of the component relationships that are found in complex environments rather than all functional interactions (Gray, 2002). In essence, part-task simulators are partial-fidelity simulators, which leave out certain characteristics in an effort to reduce the complexity of the system and to allow the users to gain proficiency in subtasks without overwhelming them with information (Stoffregen et al., 2003). Scaled worlds are effective for trainings using the part-task training (PTT) methods discussed next.

Reducing the complexity of a learning environment is often advantageous (Sweller, 1994). Research demonstrated that limiting fidelity during training can improve learning and task performance (Flach, Ficcio, McMillan, & Warren, 1986; Salas et al., 1998; Sebrechts, Lathan, Clawson, Miller, & Trepagnier, 2003; Stappers & Overbeeke, 2003; Taylor, Lintern, Koonce, Kaiser, & Morrison, 1991). Some of this advantage is achieved through the removal of perceptual clutter in the simulator, allowing focused allocation of cognitive resources to the training task. The learner can focus on the acquisition of a task component without information overload (Sweller, 1994). Also, learners may become overwhelmed and may be especially susceptible to frustration while practicing with complex, whole-task simulations (Mattoon, 1994), reducing their motivation to expend effort on the task. Thus, part-task simulators allow the partition of an otherwise difficult task into smaller, more manageable components, which is advantageous when learning involves complex tasks, such as those taught in health care.

Microworlds. Microworlds or synthetic environments (Gray, 2002) are computer simulations that mimic the component relationships found in complex systems, such as factories, forest fires, and townships. Microworlds are complex, dynamic, and opaque. Complexity refers to the number of system components and how these components interact with one another. The component relationships of microworlds can change during the simulation to correspond with actions taken by the user, making them dynamic. The opacity of a microworld indicates that internal component relationships are not visible to the user but instead must be inferred. A key advantage to the use of microworlds in the research domain is that they maintain good ecological validity and also provide a high level of experimental control (Gray, 2002). The PC-based Anesthesia Simulator developed by Anesoft is an example of a microworld (http://www.anesoft. com/products/anesthesia-simulator.aspx).

VEs. VEs are computer-generated situations with which an individual can interact (Satava, 2006; Stappers et al., 2003). They usually contain three important characteristics: a computer-generated visual image, high interactivity, and presence, providing the user with the feeling of being immersed in the simulated environment (Satava & Jones, 2003). Currently, VEs have increasing popularity in education and research. They have been used as training tools in many areas, such as spatial navigation (e.g., Sebrechts et al., 2003) and teleoperation (e.g., remote control of robotics in dangerous settings; e.g., Stappers et al., 2003). They have been used to enhance minimally invasive surgical techniques, such as laparoscopic surgery and endoscopies (Haque & Srinivasan, 2006). For example, the Virtual Environments for Surgical Training and Augmentation Project represents a research and development program with the aim of improving the understanding, assessment, and training of surgical skills using VEs for training (Tendick et al., 2000). Seymour et al. (2002) demonstrated that surgical residents whose training was complemented with the use of a virtual reality surgical simulation made six times fewer errors when performing a laparoscopic cholecystectomy. Additionally, residents who were trained with standard training methods were nine times more likely to fail to progress and five times more likely to injure the patient.

For VE research, presence and immersion are commonly used to describe the level of psychological fidelity created by physical fidelity (Witmer & Singer, 1998). Presence refers to a feeling of "being there," a suspension of disbelief of being in the simulation (Witmer & Singer, 1998), whereas immersion is the psychological state of being enveloped or absorbed by the simulation. An example of a VE in health care used for medical education and team training is 3DiTeams developed by Duke University Medical Center and Virtual Heroes, Inc. Other examples of VEs are Trauma Connect and Virtual Medical Simulation Training Center, both for training combat medics (Sotomayor, Salava, & York, 2012).

Alternative Classifications

A broader schema of describing the diversity of simulations being used in health care was developed by Gaba (2004), identifying 11 dimensions to categorize simulation applications. The dimensions Gaba (2004, p. i4) lists are: the purpose and aims of

simulation, the unit of participation in the simulation, the experience level of simulation participants, the health care domain to which simulation is being applied, the health care discipline of participating health care personal, the types of knowledge, skill, attitudes and behaviors addressed in the simulation, the age of the simulated patient, the technology applicable for simulation, the site of the simulation participation, the extent of direct participation in the simulation, and the feedback method accompanying the simulation. Gaba (2004) also discusses two of the oldest simulations used in health care: verbal role-playing and simulated standardized patients played by actors. One of the major benefits of this comprehensive approach to classify simulation applications is that it provides a standard to describe simulations. Finally, a conceptually complementary perspective on simulation is provided by Dieckmann, Gaba, and Rall (2007), who analyze and describe simulation as social practice that involves not only the simulator and other technical means but also other humans.

WHY SIMULATION IN HEALTH CARE?

In health care, inadequate training is a major contributor to high rates of error, large numbers of preventable adverse events, and many patient deaths (Kohn, Corrigan, & Donaldson, 1999; Leape, 1994; Rodriguez-Paz et al., 2009). Health care lags behind other safety critical industries, notably aviation (Berwick & Leape, 1999; Durso & Drews, 2010). Simulation training has enormous potential to improve patient safety by providing a safe training environment for repeated practice, exposure to rare, complex conditions and events, and assessment and immediate feedback on clinical performance (e.g., adherence to best practices, local practice protocols, and other measures of patient safety) (Gaba, 2004; Grenvik, Schaefer, DeVita, & Rogers, 2004; Rodriguez-Paz et al., 2009; Salas, Wilson, Burke, & Priest, 2005).

Traditional training is costly in terms of patient safety, time, efficiency, and monetary expense (Haluck & Krummel, 2000; Kneebone, 2010) and in some cases not highly effective to reinforce knowledge (Steadman et al., 2006). In addition, due to ethical and other practical reasons, patients cannot be used in many aspects of health care training (Sebrechts et al., 2003). Traditional alternate forms of training, such as the use of animals and cadavers, are often limited in their availability. Even when these alternates are available, individual variability and specimen diversity may not allow the training to be generalized (Issenberg et al., 1999; Sebrechts et al., 2003). Therefore, interactive simulations provide alternatives to traditional training because they replicate the actual clinical task and activities using training situations without the real-world risk to human health and life.

The use of simulation in health care has recently grown (Haluck & Krummel, 2000; Haque & Srinivasan, 2006) and gained wider acceptance (Issenberg et al., 1999). During the past several decades, technological advancements have had a great impact on medical diagnosis and treatment processes; however, there is a disconnect between these advances and current training practice (Haluck & Krummel, 2000). Simulation in health care has significant potential as a training and educational tool (Satava & Jones, 2003); however, as with any training technique, it requires critical

assessment to ensure effectiveness. Evaluating the effectiveness of simulators used for training comes with additional limitations. These limitations include identifying an adequate number of patients with the target condition or procedure to use as a control group, ethical concerns from involving patients at all as a control group, and small sample sizes due to a limited number of trainees (Ost et al., 2001).

Aviation and Health Care: Similarities and Differences

To increase the understanding of the role of simulation in health care, it is helpful to compare it with another domain in which simulation has been successfully used: aviation. By better understanding the similarities and differences between these domains, it is possible to identify when simulation training approaches can be adopted without modification and when a modification is needed.

Health care shares commonalities with aviation, but there are key differences between the two domains (Durso & Drews, 2010). In Table 5.1, we outline the differences between aviation and health care systems in terms of the social, natural, and technical systems and illustrate some of the limitations associated with simulation use.

Table 5.1 illustrates important differences to take into account when applying simulator training approaches from aviation to health care. Among the most important differences are those that involve issues of standardization of equipment and activities, the level of specialization involved, high inter-individual variability in patients, and differences in safety culture. Without taking into account these differences, simulation training in health care runs the risk of failure or at least the risk of being not as effective as possible. Thorough task analysis before initiating simulator training development can help avoid this risk.

Simulation in Health Care

Medical training. Many simulators used in health care have been demonstrated to be valuable and cost effective tools. One such simulator is the Harvey (Issenberg et al., 1999; see also Simulation in Health Care section), a life-sized mannequin interface simulating up to 27 cardiac conditions. Harvey allows novice students to learn basic procedures, such as recognition of heart murmurs, whereas advanced students learn how to correlate heart sounds with respiration. One of Harvey's most impressive attributes is the amount of testing performed to verify its training effectiveness. A multicenter study incorporating 208 senior medical students from five medical schools showed that 4th-year students who used the Harvey simulator as part of their cardiology elective had better post-test performance on a cardiology patient simulator than the traditionally trained control group (Woolliscroft, Calhoun, Tenhaken, & Judge, 1987). In addition, the study addressed concerns that simulator training would negatively affect physician–patient interaction because no differences were found in how patients perceived the professional behavior of the students. A second study of 203 second-year medical students demonstrated that the use of the Harvey simulator in conjunction with a physical skills course resulted in significant improvements in cardiac examination skills (Woolliscroft, Calhoun, Tenhaken & Judge, 1987). A survey completed by five medical schools using the Harvey

Table 5.1. Aviation and Health Care System Differences

Domain	Aviation	Health Care
Equipment	Standardized: Training and licensure are aircraft specific.	Lack of standardization: Wide variations in equipment; training and licensure are rarely equipment specific.
Activities and personnel	Standardized: Checklists and protocols; strong adherence.	Lack of standardization: Huge variations in local practices; weak adherence to best practices.
	Generalized: Copilot can perform the same activities as the pilot.	Highly specialized: Minimal duplication of activities across roles.
Interactions between systems	Technical: Aircraft are consistent operational environments; some nontechnical systems (human communication and weather) also present.	Socio-natural: Patients are highly variable; operational environment is inconsistent and often suboptimal; some technical systems (equipment).
	Culture: Reduced hierarchy. Individuals at different levels can directly communicate (e.g., flight attendant or copilot raising a concern with the pilot). Safety culture with anonymous reporting of near misses; proactive.	Culture: Rigid hierarchy. Minimal safety culture with few reported close calls; reactive.

simulator as part of their curriculum (Issenberg et al., 1999), reported high weekly simulator usage (22 hr) with the majority (17 hr) spent by small groups of students in self-learning mode. The remaining time involved instructors teaching.

Traditional health care training methods involve high monetary costs and large time commitments from students and instructors. The inclusion of simulators in the curriculum has the potential to provide student training comparable to traditional training while freeing up instructors. In addition, traditional training is dependent upon the condition or disease of the patient population. Time costs are involved while waiting for a patient on which to learn a procedure. Bridges and Diamond (1999) report that in the United States, nonsupply financial costs related to operating room (OR) training for general surgery residents alone is over $53 million annually. When considering other trainees (e.g., anesthesiologists, surgical trainees, nurses) within the OR and in other health care domains, the national training cost is enormous (Haluck & Krummel, 2000).

Nursing. In addition to the use of simulators to train physicians, simulators are also being used successfully to train other health care professionals, such as nurses. There is a wide range of simulators that are being used, ranging from high fidelity human patient

simulation mannequins (HPSMs) to task trainers (e.g., an arm for intravenous line insertion training). Simulation is being used in areas including fundamentals of nursing, medical, surgical, and pediatric care and high-acuity nursing. A recent survey of 1,060 pre-licensure nursing programs in the United States documented the widespread use of simulation in nursing education (Hayden, 2010). Respondents to the survey emphasized the importance of simulation in the nursing curriculum; 77% said they do or would substitute clinical time with simulation time if permitted by regulations. Given the popularity of simulation in nursing, it is important to briefly review the literature on the impact of simulation training on performance.

There is accumulating evidence that the use of simulation in nursing has a significant impact on several measures of performance. For example, Lapkin, Levett-Jones, Bellchambers, and Fernandez (2010) report in a systematic literature review that undergraduate nursing students improve in knowledge acquisition and critical thinking skills from the use of HPSMs. Similarly, Cant and Cooper (2010) report that simulation in nursing training increases knowledge, critical thinking ability, satisfaction, and confidence compared to control groups. Finally, Pauly-O'Neill (2009) demonstrated significant improvement in pediatric medication administration performance among simulation users compared to a control group.

Effectiveness of simulation training. Overall, researchers analyzing the effectiveness of health care simulators suggest that incorporating effective simulation into a traditional medical education curriculum will produce a positive outcome on student training. There is growing evidence that simulations can be useful in helping experienced health care professionals maintain and evaluate their technical skills (Haluck & Krummel, 2000; Issenberg et al., 1999). Simulators are currently used to teach procedures including neurosurgery, orthopedic surgery, plastic surgery, coronary artery bypass surgery, intraocular surgery, endoscopy and laparoscopic surgery, fluoroscopy, cardiovascular and pulmonary disease diagnosis and treatment, anesthesia, patient-centered programs, basic vital signs emergency care, infant care, neonatal care, obstetrics, phlebotomy and catheter placement, biological and chemical agent exposure diagnosis and treatment, HIV treatment, and more (Haque & Srinivasan, 2006; Issenberg et al., 1999).

In addition to training of technical skills, simulation training can also focus on the development of nontechnical skills. Examples of nontechnical skills are interpersonal skills, such as communication (Baker, Gustafson, Beaubien, Salas, & Barach, 2005), teamwork, and leadership, and cognitive skills, such as task management, situation awareness (Gaba, Howard, & Small, 1995), and decision making (see Cosby & Croskerry, 2003; Gaba, 1992). In recent years, there has been an emergence of nontechnical skill training especially in anesthesia and surgery (Fletcher et al., 2003; Gaba, 1989; Helmreich, Wilhelm, Klinect, & Merritt, 2001; Yee et al., 2005), and other medical specialties are following suit. Rall and Gaba (2005) point out the similarity of nontechnical skills required in anesthesia and those required in other areas of health care (e.g., the intensive care unit). Fletcher et al. (2003) provide an example on how to identify important nontechnical skills for anesthesiologists. The Anesthetist's Non-Technical Skills (ANTS) behavioral marker system includes four core nontechnical skills: task

management, team working, situation awareness, and decision making. Flin and Maran (2004) describe a simulator based course that targets these nontechnical skills.

Human factors–based work focusing on nontechnical skills involves different methods to identify required skills. Among those methods are simulator-based or real-life observational studies (Donchin et al., 1995), studies of cognition (Patel, Arocha, & Kaufman, 2001), survey-based methods, and root cause analyses. Ultimately, the choice of focusing on technical or nontechnical skills in simulator training is a function of the training goals (see Framework for Effective Simulation Training section). In recognition of the potential associated with the use of simulation, many institutions have implemented simulation into training programs and created simulation centers (Table 5.5 lists some examples).

Diagnostic and preoperative simulation. Simulation has great potential in the area of diagnosis. The acquisition of accurate data provides the basis for correct medical diagnoses (Satava & Jones, 2003). Data about a patient are obtained through medical history, laboratory tests, physical examination, and increasingly with the aid of diagnostic imaging techniques. Three-dimensional "simulated" representations of the human body can be created with the use of computed tomography (CT), magnetic resonance imaging (MRI), functional MRI (fMRI), positron electron tomography (PET), ultrasound, and single photon computed tomography (SPECT). Images collected for diagnostic purposes can be used to preoperatively plan procedures. The patient-specific images can be imported into a simulator to create a representation customized for that patient. This type of tailored simulation allows physicians to practice procedures with a patient-specific representation before executing them on the patient (reflecting also one of the differences between aviation and health care simulation). Applications that use such techniques include plastic surgery (J. Rosen, 1992), bony face deformities (Gleason et al., 1995), and tendon transplant in the leg (Delph & Zajac, 1992).

THE DEVELOPMENT OF EFFECTIVE SIMULATION TRAINING

For simulation training to be effective, it must support the key psychological aspects of the real-world task. In the Framework for Effective Simulation Training section, we propose a framework for understanding the psychological fidelity of a clinical task (e.g., scope of the task, the task steps, performance measures, and training goals) and using this information to choose the simulation tool. The emphasis on psychological fidelity is critical to the ultimate goal of all training: that the knowledge, skills, and abilities acquired during training transfer to improve real-world performance (Auffrey, Mirabella, Siebold, & Schaab, 2001; Baldwin & Ford, 1988; Roscoe & Williges, 1980; Salas, Tannenbaum, Kraiger, & Smith-Jentsch, 2012). Although the proposed framework does not guarantee transfer of training, it is based on the current empirical and theoretical evidence for developing effective training. In the Transfer of Training section, we discuss measuring the transfer of training.

In health care, there is growing evidence that simulation training improves clinical performance; however, there are limited data that this improved performance results in better clinical outcomes for patients (see the literature reviews and meta-analyses by Cook, Hatala, & Brydges, 2011; McGaghie, Issenberg, Cohen, Barsuk, & Wayne, 2011; McGaghie, Issenberg, Petrusa, & Scalese, 2010; Ross, Kodate, Anderson, Thomas, & Jaye, 2012; Schout, Hendrikx, Scheele, Bemelmans, & Scherpbier, 2010).

The Institute for Health Care Improvement (n.d.) recommends simulation training for high-risk situations where there is empirical evidence for its effectiveness. Empirical evidence for the effectiveness of clinical decision support systems is strikingly similar, with moderate evidence of improved clinical performance but limited data on patient outcomes (Garg et al., 2005). The general measurement challenges in health care are covered in detail in the Quality of Care and Clinical Performance section.

The lack of widespread data on training effectiveness translating into effective real-world clinical performance is not unique to health care training. Measuring transfer of training is frequently overlooked across all types of industries (Salas et al., 2012). Generally, assessment of performance in clinical practice and clinical outcomes after simulation training are uncommon (Lynagh, Burton, & Sanson-Fisher, 2007; Teteris, Fraser, Wright, & McLaughlin, 2012). Evaluating and comparing actual clinical performance across training methods in health care is especially challenging because of the variety of potential methodological and measurement confounds (see sections Transfer of Simulation Training to Clinical Practice and Quality of Care and Clinical Performance). However, simulation-based training was widely implemented in other safety critical industries (e.g., aviation, military) prior to extensive empirical evidence for transfer to real-world performance (Alessi, 2000).

Assessment of clinical performance, whether in a simulator or ideally in actual practice, is an essential component of simulation training (Issenberg et al., 1999; Issenberg & Scalese, 2008). Another aspect of assessment is to provide immediate feedback to trainees (Grenvik et al., 2004). Using simulation, trainees can be systematically evaluated with performance measures or indicators of competence (Cooper & Taqueti, 2008; Epstein, 2002; Good, 2003) for both preclinical (Fitch, 2007) and clinical training (Okuda et al., 2009). For clinical training, simulation performance has started to become part of the criteria for board certification and licensure in medicine (Dillon, 2004; Okuda et al., 2009), nursing (Nehring, 2008), and dentistry (Holmboe, Rizzolo, Sachdeva, Rosenberg, & Ziv, 2011). A common method for assessing student and trainee clinical competency is the Objective Structured Clinical Exam, which is a checklist assessment of key clinical skills administered to standardized or real patients (e.g., Harden, Stevenson, Wilson Downie, & Wilson, 1975).

A challenge to the effectiveness of simulation, whether for licensure or other purposes, is the use of one-shot testing (e.g., Chambers, Dugoni, & Paisley, 2004). Similar to the lack of widespread assessment of transfer of training, this issue is unique neither to either health care nor to simulation. There are considerable limitations with using single item measures, such as one simulation session, to assess clinical performance. Psychometric properties (i.e., measurement validity and reliability; e.g., Cronbach & Meehl, 1955) of nonrepeated, single-item measures may be inadequate, although it depends on the properties of the specific measure (Carmines & Zeller, 1979). Moreover, measurement of

clinical competency (fitness to practice) with the use of simulation and/or exams has the same psychometric challenges (Wass, Van der Vleuten, Shatzer, & Jones, 2001), specifically whether the measures generalize or transfer to actual clinical practice.

Framework for Effective Simulation Training

Effective simulation training requires replicating the psychological fidelity of the real-world task (Auffrey et al., 2001; Baldwin & Ford, 1988; Kozlowski & DeShon, 2004). Consequently, an understanding of the scope, psychological operations, performance measures, and training goals for the clinical task is needed prior to selecting the simulation tool. If the simulation tool or method is selected first, training may fit the technological capabilities but may not necessarily support the psychological fidelity of the clinical task.

To develop effective simulation training, we propose the following framework (adapted from Auffrey et al., 2001; Baldwin & Ford, 1988; Gaba, 2004; Salas et al., 2008; Salas & Burke, 2002) to guide simulation training development:

1. Procedure/task analysis: What is the procedure, the steps to successfully perform the task, and potential incidents or adverse events?

   a. Scope of the task: Does simulation training cover parts of the task or the entire task? Is the task limited to technical and procedural skills (e.g., placing sutures), diagnostic skills (e.g., selecting the appropriate sutures given the laceration, checking for the presence of foreign bodies), patient interaction (e.g., assessment of infection risk based on the cause of the laceration and patient's medical history, assessment of drug allergies), or all of the above (adpated from Hays, 2010)? Simulation training often focuses on tasks involving technical and procedural skills but need not and should not be limited to these (Gaba, 2004).

   b. Cognitive task analysis (CTA): What methods are necessary to understand task steps and sub-steps in terms of the "… knowledge, thought processes, and goal structures that underlie task performance" (Chipman, Schraagen, & Shalin, 2000, p. 3)? Numerous methods for conducting CTA, such as observation, structured and unstructured interviews, expert opinion, and instructions, exist (see Cooke, 1994). Applied CTA (ACTA) provides instructions for a practical set of methods (Militello & Hutton, 1998); Chipman et al. (2000) provide an introduction to general approaches to CTA. Overall, the CTA should incorporate the following:

      i. Evidence-based practices: Best practices for the most effective and safe delivery of treatment and care (Grol & Grimshaw, 2003; Makic, Von-Rueden, Rauen, & Chadwick, 2011; Timmermans & Berg, 2003).

      ii. Local protocols: Institution-specific instructions/guidelines that may include best practices and equipment manufacturer recommendations (e.g., Timmermans & Berg, 2003).

iii. Incidents and adverse events: A series of events or factors that could result in a patient hazard (Cook & Woods, 1994), incidents increase the likelihood of an adverse event (an injury due to delivery or failure in the delivery of treatment; Rothschild et al., 2005). Adverse events can be nonpreventable or preventable. Preventable incidents and adverse events rarely occur due to a single cause or error (see Cook & Woods, 1994; DeLucia, Ott, & Palmieri, 2009; Drews, 2011; Morrow, North, & Wickens, 2005; Reason, 1995), and the likelihood of an adverse event increases following suboptimal clinical performance and error. However, incidents and adverse events may also arise from other risk factors: equipment failure, lack of staffing, stress, and so on (also see error-producing conditions). Incorporating incidents and adverse events into simulator training promotes the transfer of knowledge and skills necessary to manage their occurrence (DeAnda & Gaba, 1990; Gaba, Howard, Fish, Smith, & Sowb, 2001; Grenvik et al., 2004). An example of an incident is a faulty ventilator circuit (Johnson, Syroid, & Drews, 2008), which becomes an adverse event if not detected and the patient stops breathing.

iv. Error-producing conditions (EPCs): Characteristics of the actual operational environment (e.g., multitasking, insufficient staff, mental and physical task demands, fatigue, and time pressure) that increase the likelihood of erroneous behavior leading to an increased risk of incidents and adverse events (Drews, 2011; Drews, Musters, & Samore, 2008). Like incidents, EPCs can be mimicked in simulation training and may help better prepare clinicians for similar conditions in the real world.

v. Equipment: Equipment used in the simulation should be identical to equipment used in the real world to facilitate matching the psychological aspects of training to the clinical task (Kozlowski & DeShon, 2004; Maran & Glavin, 2003). Thus, physical fidelity can be used to positively affect psychological fidelity.

2. Identify the training goals and performance measures

a. Training goals involve the simulation training objectives, that is, key knowledge, skills, and abilities. For the example of placing sutures, training goals are (adapted from Hays, 2010) assessment and minimization of infection risk, selection of appropriate sutures for the wound, and proper technical and procedural skills.

Specific training goals can be informed by elements in the scope of the activity and CTA. Also, goals depend upon the purpose of training, for example, preclinical, certification/licensure, or a continuing education for experienced clinicians (Dillon, 2004; Nehring, 2008; Okuda et al., 2009). One method for achievement of training goals is

meeting or exceeding a criterion-based level of performance in the simulator before advancing to the same task in the real world (Roscoe, 1991).

b. Clinical performance measures involve metrics describing clinical performance, for example, adherence to evidence-based practices and local protocols. Applied to the example of placing sutures, the aforementioned training goals can be operationalized as measures (adapted from Hays, 2010):

    i. Assessment and minimization of infection risk: Take the patient's medical history, assess the cause and severity of wound, check for presence of foreign bodies, and treat patient with antibiotics if applicable.

    ii. Selection of appropriate sutures: Assess wound severity (depth, width, and type of wound) and select sutures, including gathering required supplies.

    iii. Proper technical and procedural skills: Aseptic technique (hand hygiene, gloves, sterile equipment, no contamination of sterile equipment or gloved hands from a nonsterile surface) and suturing technique (spacing).

    iv. Error: Performance can be assessed in the context of simulated incidents, adverse events, and EPCs. If possible, performance measurements should not be limited to assessment during training but also should be measured in clinical practice. Examples include failure to ask a patient about drug allergies or administering the wrong type/dose of a medication.

c. Subjective measures, like focus groups, interviews, opinions of subject matter experts, and surveys, lack a "ground or an objective truth." Subjective metrics can provide information about the design of simulation training, perceived realism, and apparent effectiveness. Such measures can be used to subjectively assess constructs associated with performance, such as stress, cognitive workload, and fatigue. However, the use of only subjective performance measures is inadequate for evaluating the actual effectiveness of training (Bell & Waag, 1998). Furthermore, reliance on subjective measures exclusively is perilous for safety critical tasks because subjective and objective measures, although related, often diverge (Andre & Wickens, 1995).

3. Type of training: Simulation training can cover part of the task or the whole task (see scope of training). However, a combination of the task scope, CTA, training goals, and performance measures is necessary to ascertain

which specific parts of the task should be simulated or if training for the entire task is appropriate. In addition, the same information is needed to decide if it is appropriate to include incidents or mimic EPCs to help clinicians learn to prioritize important tasks.

a. Individual and team training: By reviewing the CTA, goals, performance measures, best practices, and local practices, one can ascertain if task steps require more than one person to be successfully completed. In health care, simulation training for teams has not been widely addressed (Rosen, 2008), although there are notable exceptions (e.g., Baker et al., 2005; Beaubien, 2004; Hamman, 2004; Howard, Gaba, Fish, Yang, & Sarnquist, 1992; Rosen et al., 2008; Small et al., 1999; Wayne et al., 2008). For an overview on simulation team training in health care, see Baker et al. (2005). An overview about medical team training is available in this volume (Xiao, Parker, & Manser, in press).

b. PTT and whole-task training: PTT covers critical steps or sub-steps of the activity, whereas whole-task training covers the entire activity (e.g., Wickens, Gordon, & Liu, 1998). A simulator for PTT will have lower complexity and likely lower overall physical realism, given the focus on specific critical steps, than a simulator used for whole-task training. An example of PTT is simulation of the technical and procedural skills for suturing while excluding the diagnostic assessment and clinician interaction with the patient. PTT is suited for independent components of an activity that require repeated practice to learn. Without the CTA determination of what constitutes an independent grouping of task steps, PTT development is difficult. For PTT, part-task simulators (Perkins, 2007) or scaled-world simulators (Gray, 2002) are ideal (the term *part-task simulation* originates in the medical literature, and the term *scaled world simulator* is commonly used in the psychological training literature). In both cases, only "a subset of the functional relationships found in a complex task environment" (Gray, 2002, p. 208) is selected to train specific skills. Consequentially, part-task simulators are typically relatively simple anatomical models for training specific technical and procedure skills (Perkins, 2007; Rosen, 2008). Multiple simulation tools can be combined to cover distinct tasks (Bradley, 2006); for example, training for taking a medical history could use a standardized patient as the simulation tool, whereas technical and procedural skills for suturing could be performed on a pig's foot.

c. Variable priority training (VPT): VPT uses a whole task training approach with the goal of prioritizing components of a task by controlling the allocation of attention (Wickens et al., 1998). The CTA, train-

ing goals, and other information are needed to determine if training should incorporate VPT. Training with PTT and VPT in anesthesiology leads to improved performance in the simulator (Johnson et al., 2008).

4. Simulation tool/technology: The selection of an appropriate simulation tool requires the definition of the scope of the clinical task, its psychological operations, and the training goals and performance measures. Available simulation technologies were described in Types of Simulators section.

5. Feedback method: Immediate performance feedback after task completion provides opportunities for learning (e.g., Grenvik et al., 2004), whereas structured feedback can be provided through after-action reviews (Fanning & Gaba, 2007). Feedback methods include simulator based, automatic assessment of performance in real time or by delayed performance summary; instructor assessment in real time providing feedback during simulation interruptions or simultaneously; and video-based debriefing performed at the conclusion of the simulation (Dismukes & Smith, 2001; Rudolph, Simon, Raemer, & Eppich, 2008). Feedback during simulation provides one of the best opportunities for learning since in clinical practice, immediate feedback is rare (Durso & Drews, 2010).

Other factors relevant to the effectiveness of training are not explicitly covered in this framework. Among these are trainee characteristics, such as background knowledge, ability, and motivation (Baldwin & Ford, 1988). Participant motivation can be bolstered by providing a clear purpose and set of expectations for training (Baldwin & Ford, 1988) and by presenting training as an opportunity instead of an evaluation (Salas et al., 2012). Incorporating progressively increasing difficulty levels and repeated practice are recommended for effective training (Issenberg & Scalese, 2008). Finally, organizational factors may impede or impair the success of simulation training. Salas et al. (2012) put forward a comprehensive approach prior to developing training called a training needs analysis to determine "expected learning outcomes, guidance for training design and delivery, ideas for training evaluation, and information about the organizational factors that will likely facilitate or hinder training effectiveness" (p. 80).

Overall, the aforementioned framework is conceptual and provides guidance to help identify the psychological operations for the clinical task that guide the selection of the simulation technology. However, its application does not always guarantee positive transfer since the amount of psychological correspondence necessary for positive transfer is not known (other than the generic statement of as much as possible) and must be inferred (see Inferring Transfer of Training).

Training myths. In addition to factors that increase the likelihood of training effectiveness, there are intuitively appealing but misguided factors that may decrease training effectiveness. Two myths, cognitive learning styles and high-(physical)-fidelity simulation, are dispelled.

The pervasive belief that instruction is most effective when it conforms to an individual's cognitive or learning style, such as kinesthetic, auditory, or visual, is a myth (Pashler & McDaniel, 2008; Riener & Willingham, 2010). The research and applications of learning styles often focus on teaching children (e.g., Riener & Willingham, 2010) but have also been suggested for tailoring simulation training to the preferred learning styles of clinicians (e.g., Newble & Entwistle, 2009; Weaver et al., 2010). Unfortunately, there is almost no empirical evidence that cognitive/learning styles should be taken in account because, regardless of individual preferences, information presented in multiple modalities consistently leads to the best acquisition of knowledge (Pashler & McDaniel, 2008; Riener & Willingham, 2010). Given the absence of empirical support for learning styles in the education and psychology literature, it is unlikely that tailoring the modalities of instruction, such as the simulation tool, to preferred learning styles will increase training effectiveness. Instead, we recommend tailoring simulation to the training goals, specifically, psychological fidelity, which is empirically supported.

The second myth is the belief that high-fidelity simulation (i.e., strong physical realism of the simulation tool and environment) is by itself sufficient to make training effective. If the physical realism in the simulation does not correspond to the psychological aspects of performing the actual task, high physical fidelity even may be detrimental to training (Kozlowski & DeShon, 2004). The erroneous emphasis on high-fidelity simulation haunted aviation for decades, increasing the cost of simulation, limiting its availability, and perhaps even diminishing its effectiveness (see Alessi, 2000; Hopkins, 1974).

Transfer of Training

The most important objective of training is successful transfer; that is, the knowledge, skills, and abilities acquired during training translate into improved real-world performance (Baldwin & Ford, 1988; Salas et al., 1998, 2012). Transfer of training is suggested by improved performance in the simulator, but that improvement is not the definitive indicator. To validate simulation performance evaluations, a comparison with real-world performance is required (Devitt, Kurrek, Cohen, & Cleave-Hogg, 2001; Dong et al., 2010; Johnson, Guediri, Kilkenny, & Clough, 2011; Schout et al., 2010). Training that leads to increased real-world performance is a result of positive transfer (Osgood, 1949), and high psychological fidelity is central for positive transfer to occur (Adams, 1979; Kozlowski & DeShon, 2004). When training neither increases nor decreases real-world performance, there is zero transfer. Negative transfer (Osgood, 1949) involves training that decreases real-world performance and results from an incompatibility in applying the knowledge, skills, and abilities acquired during training to the real world (Lui, Blickensderfer, Macchiarella, & Vincenzi, 2008). Negative transfer can also result from mismatched or incorrect psychological fidelity, for example, learning the incorrect sequence of steps, irrelevant tasks, distracting high physical fidelity, or equipment differences, which results in habit interference (Lui et al., 2008).

The concept of transfer of training represents a continuum (Lui et al., 2008). Positive transfer from simulation training to the real world is often partial, especially during the early stages of training. Some but not all skills and knowledge successfully transfer from the simulation to the real world, and the rates of learning may be lower than real-world training (Schmidt-Panos & Scerbo, 2008). For many tasks in health care, one of the primary goals of simulation training is not just the initial acquisition of knowledge and skills but also their retention. Although simulation has been proposed to improve skill retention (Ziv et al., 2006), longitudinal assessment of performance in health care is uncommon (although see Wayne et al., 2006).

Transfer of simulation training to clinical practice. Literature reviews and meta-analyses indicate that simulation training in health care is highly effective (McGaghie et al., 2010, 2011; Ross et al., 2012). However, the majority of reviewed studies used performance measures limited to the simulator and failed to assess actual clinical practice. Still, positive transfer of simulation training has been found in a variety of activities, shown in Table 5.2.

The limited empirical assessment of transfer of training in the health care setting reflects various challenges associated with measuring real-world clinical performance. Measuring human performance can be expensive and time consuming, require institutional support, and have methodological confounds (e.g., uncontrollable factors, such as patient health and the unpredictability of responses to treatment; Gaba, 2004; Schout et al., 2010). Another challenge arises from ethical and legal considerations of continued use of traditional training methods with more effective methods being available (Ziv et al., 2006).

Theories of transfer of training. Theories of transfer of training can be contrasted from two positions (Lathan, Tracey, Sebrechts, Clawson, & Higgins, 2002; Singley & Anderson, 1989):

1. General transfer: Transfer is broad and generalizes from one task to another, even when unrelated ("Doctrine of Formal Discipline"; Angell, 1908, as cited in Lathan et al., 2002). For example, the knowledge, skills, and abilities acquired from simulation-based training for suturing a cut on a pig's foot generalizes to suturing cuts, in a variety of locations, on patients.
2. Specific transfer: The two tasks must be identical for transfer to occur ("Theory of Identical Elements"; Thorndike, 1906, as cited in Lathan et al., 2002). An example of specific transfer is that simulation training for suturing using a pig's foot transfers only to the exact same task and simulation tool, that is, suturing using a pig's foot.

Both theories have flaws: Evidence for general transfer is scarce because skill generalization between tasks with unrelated psychological operations is difficult to establish (Singley & Anderson, 1989). Specific transfer implies that it is necessary to train under

Table 5.2. Examples of Positive Transfer of Simulation Training

Tasks	References
Anesthesiology: Intubation and ventilation techniques, correct application of pressure to prevent regurgitation during anesthesia, and management of incidents and adverse events.	DeAnda & Gaba (1990); Domuracki, Moule, Owen, Kostandoff, & Plummer (2009); Gaba, Howard, Fish, Smith, & Sowb (2001); Ross, Kodate, Anderson, Thomas, & Jaye (2012).
Bronchoscopy: Diagnostic viewing of the airways and the lungs.	Ross et al. (2012).
Cardiac arrests (resuscitation when the heart stops beating).	Edelson et al. (2008); Wayne et al. (2008).
Central line placement: A catheter inserted near the heart, used to infuse medications and draw blood.	Barsuk, Cohen, Feinglass, McGaghie, & Wayne (2009); Barsuk, McGaghie, Cohen, Balachandran, & Wayne (2009); Barsuk, McGaghie, Cohen, O'Leary, & Wayne (2009); Dong et al. (2010); Evans & Dodge (2010).
Diagnosis of heart murmurs.	Fraser et al. (2011).
Dystocia: A complication in birth in which the baby's shoulder blocks delivery.	Draycott, Crofts, & Ash (2008).
Surgical procedures: Colonoscopy; robot-assisted techniques, such as laparoscopic surgery; etc.	Al-Kadi et al. (2012); Park et al. (2007); Schout, Hendrikx, Scheele, Bemelmans, & Scherpbier (2010); Seymour (2008).

every possible condition that could occur in the real world, which is impossible (Singley & Anderson, 1989). Furthermore, specific transfer requires high physical fidelity, a requirement that is at odds with research demonstrating the effectiveness and importance of psychological fidelity.

The theory of transfer appropriate processing (Morris, Bransford, & Franks, 1977) falls between the general transfer and specific transfer extremes. It predicts that when the initial task (i.e., simulation training) supports the psychological operations of the second task (i.e., clinical practice in the real world), then positive transfer occurs. In order to develop training that is likely to transfer to a reasonable range of real-world conditions, skills can be practiced repeatedly with performance feedback until they approach automaticity; separate training can be conducted for independent, critical steps using PTT; and a final training can be conducted under conditions that mimic the operational environment (i.e., simulated incidents and adverse events and EPCs). Compared to simulation training with transfer appropriate processing, traditional training does not typically permit repeated practice under safe and standardized conditions.

Psychological fidelity and transfer. An important lesson in developing effective simulation training in health care can be taken from aviation. For decades, aviation training singularly focused on high-fidelity simulation and failed to produce tangible benefits (Alessi, 2000). Questions that were raised about the significant expenses and lack of

benefit from aviation simulators with "bells and whistles" (Hopkins, 1974) went unheeded. Empirical work on the transfer of training in aviation repeatedly demonstrated the importance of psychological fidelity across numerous tasks (Adams, 1979; Hays, Jacobs, Prince, & Salas, 1992; Kozlowski & DeShon, 2004). Only when high physical fidelity also has psychological relevance to the task has it been shown to improve transfer of training, for example, haptic and visual feedback for laparoscopic surgery (Hays et al., 1992). For simulation training to become widespread in health care, it needs to be cost effective. For that reason, the same pursuit of naive realism as in aviation needs to be avoided. Fortunately, more and more researchers in health care identify the importance of psychological fidelity as a critical dimension for effective training (Beaubien, 2004; Maran & Glavin, 2003; Scerbo & Dawson, 2007). Still, the allure of naive realism in simulation is enduring (Hook, 2004; Smallman & St. John, 2005).

Quality of Care and Clinical Performance

The overarching goal of simulation training is the safe and effective delivery of high quality care. In contrast to clinical performance, quality of care is a long-term, multidimensional metric that requires multiple measures, including mortality, clinical outcomes, patient functional status, and patient quality of life (see Clancy, 1998). Because a multitude of factors impact quality of care and clinical outcomes, it is difficult to determine if simulation training improves performance on all of these measures.

The goal of simulation training is improved quality of care and clinical outcomes, but each is difficult to assess. Hence, we contend that a more easily measured, immediate, and directly task-relevant level of analysis is assessment of clinical performance. Clinical performance is a predictor of risk for suboptimal clinical outcomes and thus quality of care, just as "near misses" and "close calls" in aviation are risk predictors for accidents and fatalities. In health care, "near misses" can be inferred using clinical performance measures, such as nonadherence to best practices and error (Drews, in press).

To conceptualize and understand the relationships between quality of care, clinical outcomes, and clinical performance, we use Simon's (1996) theory of complex, hierarchical systems. In a hierarchical system, there are interrelated systems with each level influenced by all of the prior levels (Simon, 1996). As the level of analysis increases, the number of related systems and subsystems grows, increasing the complexity. Here, the highest level of analysis is quality of care (incorporating clinical outcomes and clinical performance), followed by the midlevel of clinical outcomes (including clinical performance), and the lowest level of clinical performance (Table 5.3).

In health care, the time between the levels of analysis varies by orders of magnitude. The majority of clinicians do not receive immediate feedback on their clinical performance, which impacts clinical outcomes and influences the quality of care. The epochs and open-loop feedback in health care can be contrasted with aviation, whereby pilots generally receive immediate or close-to-immediate feedback on the majority of their actions, that is, closed-loop feedback.

Clinical outcomes and quality of care are influenced by multiple tasks or procedures and other factors, such as patient health. In addition, there is a limited subset of tasks

Table 5.3. Clinical Performance: Hierarchical Levels of Analysis

Level of Analysis	*Dependent Variables*	*Time Scale*
Quality of care	A multidimensional measure that takes into account clinical outcomes, mortality rates, and patient reported data, such as functional status and quality of life.	Typically weeks, months, and years. In emergency care and other specialized areas, such as critical care or anesthesiology, there may be a more direct relationship to lower levels, with clinical performance affecting clinical outcomes immediately.
Clinical outcomes and adverse events	Medical and laboratory record abstraction of physiological and diagnostic tests, for example, infections. Clinical outcomes and adverse events can be derived from objective classification and subjective classification by expert and nonexpert clinicians.	Days, weeks, and longer.
Clinical performance	Observations or video recording of adherence to best practices, local practice protocols, incidents, and error. May also be inferred from item and equipment use.	Seconds, minutes, and hours.

Note: Hierarchical levels of analysis for clinical performance, dependent measures, and time scale. The highest level of analysis is quality of care, followed by clinical outcomes and, last, clinical performance. Over time, each level integrates all elements of the lower levels. That is, quality of care is associated with clinical outcomes and clinical performance because adherence to best practices increases the likelihood of positive clinical outcomes, which are indicators and causal factors for high quality of care.

and activities in health care with immediate consequences. Exceptions include specific situations in the areas of emergency care, critical care, respiratory therapy, surgery, and anesthesiology. In these areas, clinical outcomes and mortality may be on the time scale of seconds, minutes, or hours.

To limit the scope of the hierarchical system, other potentially relevant factors (e.g., the operational environment and social/cultural factors) are excluded. Clinical performance could be further reduced to the impact of adherence or nonadherence to best practices. In health care, the relationship between time and levels of analysis is more formally described using the two propositions for nearly decomposable systems (Simon, 1996):

1. Over a short time scale, levels are mostly independent of each other.

2. Over a long time frame, the outcome at the higher levels depends upon the aggregation of the outcomes of the levels below.

Clinical Performance

The hierarchical level of analysis is a theoretical framework that explains how clinical performance acts as a precursor to clinical outcomes and how clinical outcomes impact quality of care. Clinical performance is the common level of analysis for both simulation training and the real world. Methodological considerations for simulation training include skill retention, type of training (e.g., repeated practice and incidents, VPT, and PTT), and clinical performance measures.

For simulation training to be effective, it must both transfer to the real world, and the knowledge, skills, and abilities acquired in training must be retained over time. Practice is fundamental to achieving expert performance (Ericsson, Krampe, & Tesch-romer, 1993). Achieving expert performance may require 10,000 hr of practice (Simon, 1996).

Several theoretical frameworks on skill acquisition (e.g., Fitts, 1964; Rasmussen, 1986) focus on the importance of practice, not necessarily transfer. Nonetheless, these frameworks provide a solid foundation for skill acquisition and retention that is missing from the transfer theories discussed earlier. All of these frameworks propose a three-stage learning process with stages consisting of declarative knowledge (learning stage), procedural knowledge (consolidation stage), and a combination of declarative and procedural knowledge (tuning stage) (Kim, Ritter, & Koubek, 2013). The third stage of learning constitutes expert performance. Kim et al. (2013) propose that each stage of learning has different properties for forgetting. Because knowledge in the first stage of learning is strictly declarative, there is potential for memory failures, with much of the learning from practice being lost. With sufficient practice, knowledge becomes increasingly procedural and thus more resilient to forgetting.

Another issue in training is the spacing between learning, that is, should it be massed (repeated back-to-back training) or distributed practice (more time between training intervals) (Cepeda, Paschler, Vul, Wixted, & Rohrer, 2006)? The type of task, goals, and performance measures should be used to determine whether practice is massed or distributed. Massed practice may lead to better acquisition and retention for skills that are primarily procedural (Cepeda et al., 2006), such as suturing. For retention of declarative knowledge, distributed practice is generally superior to massed practice. However, the length of the spacing intervals is uncertain (Cepeda et al., 2006). Because simulation training may consist of a task with procedural and declarative components, it may be advisable to separate the task components using PTT. Massed practice on procedural and declarative components can be trained on a distributed schedule.

Based on the theories of learning, skill acquisition, and forgetting, simulation training should reach at least the second stage of learning and realize a minimum level of proficiency and reduced skill decay to prevent memory failures. To determine if the second stage of learning is reached, simulation practice requires repetition until a minimum performance criterion is reached. However, it is unlikely that simulation

training will ever be extensive enough to achieve expert performance, the third stage of learning.

Measurement Methods

Multiple methods and measures for evaluating clinical performance exist, such as video recording, observation, metrics provided by the simulation tool (e.g., chest compression force for administering cardiopulmonary resuscitation on a mannequin), and inventory records (for item use). Next, the advantages and disadvantages of different methods of measuring clinical performance are discussed.

An advantage of video recording is that performance can be coded by multiple raters; thus psychometric properties of measures, such as inter-rater reliability, can be quantified and reviews of videos can be used to provide trainees with explicit performance feedback. Another advantage is that coding can be performed using computer vision algorithms, which may reduce the need for human raters, may lower costs, and may even provide automated, structured performance feedback (e.g., Zhang & Li, 2010). Although video recording is easily employed in simulation training, in a clinical setting, concerns of patient privacy and potential legal liability complicate its use. Observation is typically the most common method of measurement.

With observation, performance must be coded in real time; thus if clinician actions are missed by the coder, they cannot be reviewed. Because there is no video recording, observation cannot provide the same level of detail in the performance feedback. Observation and video coding have the same drawback: requiring a rater, or ideally multiple raters, with sufficient clinical knowledge and expertise to conduct valid and reliable coding.

In actual clinical practice, item use provides a simple, albeit approximate, measure of clinical performance. Item use can be informative for tasks that occur with regularity (e.g., oral care and central line maintenance). When the use of items is indicative of adherence to best practices, lack of item consumption allows potential identification of nonadherence. However, knowing a task was performed is not sufficient to determine if best practice standards were followed. In addition, item consumption relies on the accuracy and capabilities of the inventory system. If the health care facility has an inventory system linking items to specific patients, then inferences at the patient level are possible.

Operationalizing clinical performance. Clinical performance measures can be operationalized using adherence to best practices and local practice protocols, the frequency and type of error, item use, and self-report measures. Different clinical performance measures may overlap. For example, some best practices may be included in local practice protocols, and deviations from evidence-based practices can be considered as error.

Specific metrics of clinical performance should be selected using evidence based practices, local practice protocols, training goals, and the results of the CTA. To have a well-defined relationship between clinical performance and outcomes, best practices that are empirically supported should be used. There are wide variations in the strength of the evidence for best practices. These levels range from well-established empirical

evidence, followed by expert opinion and manufacturer guidelines or recommendations (Makic et al., 2011). If "incidents" are part of simulation training, there may be specific best practices for particular situations, which may also depend on the local practices, training goals, and the CTA. Incidents may be rare in clinical performance (although see Medmarx for a medication error database; http://www.medmarx.com), so measurement of transfer performance for incidents may not be possible.

Error may consist of best practice nonadherence but may also reflect omission of procedure steps, forgetting items, and other actions that are implied in best practices. There are causal descriptions of error, which are more detailed than adherence or non-adherence to a standard. Error can be classified using the Reason (1990) taxonomy (see also Drews, 2011; Drews, Wallace, Benuzillo, Markewitz, & Samore, 2012; Morrow et al., 2005):

1. Slip: Breakdowns in selecting or recognizing appropriate actions
2. Lapse: Memory or attention failures
3. Mistake: Incorrect choices or wrong objectives
4. Violation: Conscious avoidance of normal practices, that is, intentional "shortcuts."

Inferring transfer of training. From a methodological perspective, an ideal research design directly compares real-world performance from simulation training to that from traditional training (see Strayer & Drews, 2003). However, this type of research design may not be possible due to considerations outlined earlier. Fortunately, there are several alternatives to infer transfer of training.

Simulation validation. Clinical performance measures for the simulation can be validated by comparing clinical performance in the simulation and clinical performance in the real world (e.g., Schout et al., 2010). There are numerous possible confounds, for example, inconsistencies between real patients and the simulation tool. Nevertheless, comparable performance in simulation and clinical practice implies that simulation-based performance assessment is a valid proxy for real clinical practice. Furthermore, validation makes it possible to control for traditional training by evaluating performance in the simulator.

A complementary approach toward validation is to compare simulator performance of participants with different levels of expertise. Demonstration of group differentiation based on skill levels can serve as a validation of the simulated scenarios (Murray et al., 2007).

Quasi-transfer. Related to validation of simulation performance, quasi-transfer is a method for testing the generalizability of performance by using one simulation method for training and a second one for transfer. An example of quasi-transfer is using virtual reality for surgery training with surgery on human cadavers to evaluate transfer (e.g., Schout et al., 2010). In aviation, quasi-transfer has been demonstrated to be predictive of real-world performance (Taylor et al., 1993).

Training effectiveness. There is also a set of descriptive methods for comparing training effectiveness in terms of time or cost using ratios, that is, percentage transfer, transfer effectiveness, and incremental training effectiveness (Roscoe & Williges, 1980). Developed for aviation, these methods allow quantifying the time and costs of simulation versus real world in terms of training effectiveness. However, in health care, ethical considerations and the quantification of the expenses of clinician "learning" on patients is a challenge to this approach.

Clinical outcomes. Clinical outcomes can be measured via abstraction of medical records, including laboratory, physiological, and diagnostic tests. Some clinical outcomes are rare, so years of patient data may be required to determine the effectiveness of simulation training for a single institution. Past data on clinical outcomes can be used to compare with simulation training data, but such simple pre- and post-training research designs have potential for numerous confounds, such as improved equipment, increased awareness of best practices due to continuing education, and changes in local practices, protocols, and staffing, to name a few.

Objective, systematic methods should be used to select clinical outcome measures (see Campbell, Braspenning, Hutchinson, & Marshall, 2002). Any subjective classification of clinical outcomes should be interpreted with caution because the frequency and severity of negative clinical outcomes may be underestimated. For example, Lin et al. (2010) found that subjective classification of CLABSI rates in hospitals were as much as four times lower than objective classification. Also, abstracted information from medical records reviewed by expert raters was inconsistent and tended to systematically underreport the severity of diseases and illness (Luck, Peabody, & Dresselhaus, 2000).

Similar to clinical performance, objective methods for determining clinical outcomes is essential to the validity and reliability of the measurement construct. Comparisons in the rates of negative clinical outcomes between health care facilities should be interpreted carefully (Luck et al., 2000). Hospitals that frequently perform high-risk procedures are more likely to have sicker patients and higher rates of negative clinical outcomes compared to hospitals that perform fewer or none of the same high-risk procedures. Despite the limitations of measuring clinical outcomes, they are an essential component in a cost analysis of simulation training.

COST ANALYSIS OF SIMULATION TRAINING

Compelling data exist indicating that simulation training can save money compared with traditional training. One possible reason for the low adoption rate of simulation training is its perceived high cost. In addition to the monetary expense of the simulation tool, other expenses include the costs of training development, staff and instructor time, and evaluation. Few studies report the costs of simulation training (Zendejas, Wang, Brydges, Hamstra, & Cook, 2013). Simulation systems can range in cost from less than $5,000 for laparoscopic simulators to over $300,000 for sophisticated anesthesia simulators (Issenberg et al., 1999). Student trainees increase patient safety risks.

Although traditional training protocols involve the guidance of an experienced mentor (Haque & Srinivasan, 2006), the trainee may not always follow the protocol (Haluck & Krummel, 2000).

In high risk areas of health care (e.g., anesthesiology, critical care, emergency care, obstetrics, and surgery), simulation training is associated with reduced negative clinical outcomes, which are linked with malpractice claims (Hanscom, 2008). For central line placement, a high-risk procedure, Cohen et al. (2010) analyzed costs and benefits for simulation training and found the benefits were far greater than the costs: For each dollar spent on training, $7 was saved in direct care costs (Table 5.4).

Two related methods to compare costs and benefits are CBA and return on investment (ROI) (e.g., Feldstein, 2011). In a CBA, total benefits are subtracted from total costs. If the CBA is positive, simulation training results in cost savings. For the ROI, the "return" (total benefits) is divided by each dollar "invested" (total costs) on training. If the ROI is greater than one, simulation training produces cost savings. A positive CBA implies a positive ROI and vice versa.

Cost values from the CBA and ROI performed by Cohen et al. (2010) are shown in Table 5.4, comparing the costs and benefits for simulation training for central line placement and the hospital expenses for a CLABSI, a preventable, life-threatening, negative clinical outcome.

The previous example does not include uncertainty in the cost analysis, which can be taken into account using sensitivity analysis (e.g., Drummond, Sculpher, Torrance, O'Brien, & Stoddart, 2005). A sensitivity analysis could be incorporated by calculating the minimum and maximum cost for extra hospital days based on less expensive care in a step-down unit and more expensive care in a critical care unit.

Types of costs: Fixed, variable, direct, and indirect. Costs can be categorized as either fixed or variable and direct or indirect (e.g., Feldstein, 2011). Fixed costs are one-time expenses, which in the previous example are the ultrasound machine and the central line simulator. Variable costs are ongoing expenses and in simulation training depend on the number of trainees and other factors, for example, number of kits and gowns, simulation facility charges, and amount of salary support for instructors and staff. Total costs are the sum of all fixed costs and all variable costs (Feldstein, 2011).

Table 5.4 excludes indirect costs. From the perspective of the health care facility, direct costs are delivery of care expenses (Scott, 2009). Direct costs of negative clinical outcomes can be obtained from published research and internal billing records. There is considerable uncertainty in estimation of costs of negative clinical outcomes. For example, Scott (2009) used a sensitivity analysis to estimate the 2007 direct hospital costs for CLABSI, which range from $670 million to $2.68 billion in the United States. For the health care facility, examples of indirect costs are losses because of decreased rates of insurance reimbursement due to frequent adverse clinical outcomes, such as health care associated infections, and the costs associated with increased malpractice claims. For the patient, indirect costs include lost income, diminished quality of life, and mortality (Scott, 2009). Indirect costs are more difficult to measure and estimate than direct costs because they have a greater uncertainty (see Andel, Davidow, Hollander, & Moreno, 2012).

Table 5.4. Cost Analysis of Simulation Training for Central Line Insertions

Costs	Amount (US$)
Equipment (e.g., ultrasound, central line simulator, kits, gowns, and other items)	$39,294
Facility rental	$14,850
Salaries: Instructor and research assistant	$57,727
Total training costs	$111,916
Benefits (cost savings)	
Simulation training resulted in 137 fewer patient hospital days due to a lower frequency of central line associated bloodstream infections (9.95 estimated infections with a predicted average of 13.77 additional hospital days per infection)	$823,164
Cost-benefit analysis (CBA)	
Benefits ($823,164) – Total Costs ($111,916) The total benefits minus the total costs.	$711,248
Return on investment	
CBA / Training Costs = $711,248 / $111,916 The return for each dollar spent on training.	7 to 1

Note: This example is adapted from Cohen et al. (2010), with simplified values. The table shows simulation training costs, benefits, the cost-benefit analysis for training, and the return on investment for training.

Limitations

As with any other quantitative method, cost analysis depends upon the reliability and validity of the input values. The uncertainty of estimating costs and benefits is compounded by the need to objectively measure clinical outcomes (Lin et al., 2010). Proxy measures can be used to infer risk for negative clinical outcomes, such as suboptimal clinical performance and item or equipment usage. To conduct a comprehensive cost analysis that includes risk, probability distributions can be assigned to values and modeled (Drummond et al., 2005). Estimates for indirect costs are substantial, possibly in the hundreds of billions of dollars in the United States, but are often overlooked because they are difficult to quantify (Andel et al., 2012).

Comprehensive reviews of cost analysis methods can be found in Feldstein (2011) and Drummond et al. (2005); Andel et al. (2012) present recent data on different types of costs estimates for medical errors.

CHALLENGES AND THE FUTURE OF SIMULATION

Preventable adverse clinical outcomes are estimated to result in 98,000 lives lost every year and a projected cost between $17 billion and $29 billion in the United States (Kohn, Corrigan, & Donaldson, 1999); newer cost analyses are consistent with

this range and indicate total costs (including adjusted quality of life) may reach a staggering $735 billion to $980 billion dollars (Andel et al., 2012). There is an obvious need to improve patient safety, and simulation is one important tool in this pursuit. However, numerous challenges prevent health care from embracing simulation training as one of the solutions.

Among these challenges is the strong belief that the development of effective simulation training is solely an engineering problem rather than a human factors and ergonomics challenge. This sentiment finds its expression in the perceived importance of naive realism in the context of simulator development. Human factors professionals encounter this sentiment frequently. It is up to our profession to help developers of health care simulators to move beyond the fascination with possible technological solutions, without understanding the problem, to the implementation of solving real problems using technology. As such, it will be one of our field's main tasks to emphasize the importance of psychological fidelity rather than physical fidelity for simulation to succeed.

Another challenge is that simulator curriculum development can be based on traditional clinical curricula, implying very little need to perform precise analytical work preceding the development and implementation of simulator-based training scenarios. For simulation to succeed in health care, it will be critical to perform careful and comprehensive task analyses that are informed by the human factors and ergonomics literature. Providing the simulation training community with the results of task analyses in a range of specialties will significantly increase the speed of adoption and success of simulation for training. In addition, the results of task analytical work will also improve our understanding of human performance, breakdowns of performance, and the conditions that need to be present to minimize error and nonadherence to best practice.

An additional challenge in this context is that there is very little effort devoted to the development of a comprehensive library of standard scenarios that can be used for health care workers in the different specialties. Two exceptions for standard scenarios include the website for the Society for Academic Emergency Medicine Simulation Case Library and publications of simulation scenarios in the journal *Simulation in Healthcare* (see Table 5.5). Currently, there are many simulation centers nationwide that develop their own scenarios to train health care workers, which leads to repetition of development work and may result in sub-optimal simulation trainings because of resource limitations. Development of a national library of standard scenarios for use in simulators is highly desirable in health care and can contribute significantly to wider adoption and more successful simulator training in health care.

As technology advances over the next decades, so will the sophistication and use of simulation training in health care. The number of simulation centers in hospitals across the United States and the amount of resources available for use of simulation training will increase, and there will be a move toward the use of simulation as one important step toward an assessment of knowledge, skills, and abilities of clinicians. Given the success of simulation training in aviation, repeated use of simulation training and simulation based certification of health care providers is likely a critical step toward improvement of quality of care and patient safety.

Finally, with the advance of personalized medicine, the use of simulation to explore different procedural approaches and techniques even before a procedure is

Table 5.5. Recommended Resources

Category	Resource	Reference
Academic books and journals	Institute of Medicine report on patient safety	Kohn, Corrigan, & Donaldson (1999)
	"Error in Medicine" (classic article on error in medicine)	Leape (1994)
	Human Factors (journal)	http://hfs.sagepub.com/
	Simulation in Healthcare (journal)	http://journals.lww.com/simulationinhealthcare/pages/default.aspx
	"The Economics of Health Care Quality and Medical Errors" (recent paper on direct and indirect cost estimates for medical errors)	Andel, Davidow, Hollander, & Moreno (2012)
Fun reading	*Complications: A Surgeon's Notes on an Imperfect Science* (book)	Gawande (2007)
Other resources	Agency for Healthcare Research and Quality	http://www.ahrq.gov/
	Training Guide: Using Simulation in TeamSTEPPS Training	http://www.ahrq.gov/teamsteppstools/simulation/index.html
	Institute for Healthcare Improvement	http://www.ihi.org
	Simulation for High-Risk Situations	http://app.ihi.org/imap/tool/#Process=f7a57213-3446-4539-befb-66f8269c67fb
	Society for Academic Emergency Medicine (SAEM) Simulation Case Library	http://www.emedu.org/simlibrary/
Simulation centers	Advanced Simulation Clinic, School of Dentistry, University of Minnesota (tour)	http://sodfiles.ahc.umn.edu/ASC_Orientation/student.html
	Center for Immersive and Simulation-Based Learning, School of Medicine, Stanford University	http://cisl.stanford.edu
	Intermountain Healthcare Simulation Learning Center, College of Nursing, University of Utah	http://nursing.utah.edu/simulation-learning-center/index.php
	Johns Hopkins Simulation Center	http://www.hopkinsmedicine.org/simulation_center/
	Michael S. Gordon Center for Research in Medical Education, Miller School of Medicine, University of Miami	http://www.gcrme.med.miami.edu/
	University of California–Los Angeles Simulation Center	http://www.sim.ucla.edu/

Table 5.5. (continued)

Category	Resource	Reference
	The Center for Medical Simulation, Harvard University	http://www.harvardmedsim.org/
Societies	Human Factors and Ergonomics Society	https://www.hfes.org/web/Default.asp
	International Nursing Association for Clinical Simulation and Learning	https://inacsl.org/
	Society in Europe for Simulation Applied to Medicine	http://www.sesam-web.org/
	Society for Simulation in Health Care	http://ssih.org/

being performed provides a great opportunity to increase the quality of care for patients in general and for critically ill high-risk patients in particular.

At this point, there is an opportunity to accelerate the development and adoption of effective simulation training in health care by including human factors and ergonomics as a significant contributor. Human factors and ergonomics can provide the theoretical basis for simulation in health care and, as such, allow leveraging one century of experience accumulated in aviation simulation training. For human factors and ergonomics, this experience provides a great opportunity to help accelerate advancement in patient safety through the development of effective simulation training.

RECOMMENDED RESOURCES

Table 5.5 contains recommended resources for additional reading, such as books and journals, websites simulation centers, and societies.

ACKNOWLEDGMENTS

REFERENCES

Abrahamson, S., Denson, J. S., & Wolf, R. M. (1969). Effectiveness of a simulator in training anesthesiology residents. *Journal of Medical Education*, *44*, 515–519.

Adams, J. A. (1979). On the evaluation of training devices. *Human Factors*, *21*, 711–720.

Alessi, S. M. (2000). Simulation design for training and assessment. In H. F. O'Neil, & D. H. Andrews (Eds.), *Aircrew training and assessment* (pp. 199–224). Mahwah, NJ: Lawrence Erlbaum.

Al-Kadi, A. S., Donnon, T., Oddone Paolucci, E., Mitchell, P., Debru, E., & Church, N. (2012). The effect of simulation in improving students' performance in laparoscopic surgery: A meta-analysis. *Surgical Endoscopy*, *26*, 3215–3224.

Andel, C., Davidow, S. L., Hollander, M., & Moreno, D. A. (2012). The economics of health care quality and medical Errors. *Journal of Health Care Finance*, *39*, 39–50.

Andre, A. D., & Wickens, C. D. (1995). When users want what's not best for them. *Ergonomics in Design*, *3*(4), 10–14.

Auffrey, A. L., Mirabella, A., Siebold, G. L., & Schaab, B. (2001). *Transfer of training revisited*. Alexandria, VA: Army Research Institute for the Behavioral and Social Sciences.

Baker, D. P., Gustafson, S., Beaubien, J. M., Salas, E., & Barach, P. (2005). Medical team training programs in health care. In K. Henriksen, J. B. Battles, E. S. Marks, & D. I. Lewin (Eds.), *Advances in patient safety: From research to implementation. Vol. 4: Programs, tools, and products* (AHRQ Publication Nos. 050021). Rockville, MD: Agency for Healthcare Research and Quality.

Baldwin, T., & Ford, J. (1988). Transfer of training: A review and directions for future research. *Personnel Psychology*, *41*, 63–105.

Barnes, R. W., Lang, N. P., & Whiteside, M. F. (1989). Halstedian technique revisited. Innovations in teaching surgical skills. *Annals of Surgery*, *210*, 118–121.

Barsuk, J. H., Cohen, E. R., Feinglass, J., McGaghie, W. C., & Wayne, D. B. (2009). Use of simulation-based education to reduce catheter-related bloodstream infections. *Archives of Internal Medicine*, *169*, 1420–1423.

Barsuk, J. H., McGaghie, W. C., Cohen, E. R., Balachandran, J. S., & Wayne, D. B. (2009). Use of simulation-based mastery learning to improve the quality of central venous catheter placement in a medical intensive care unit. *Journal of Hospital Medicine*, *4*, 397–403.

Barsuk, J. H., McGaghie, W. C., Cohen, E. R., O'Leary, K. J., & Wayne, D. B. (2009). Simulation-based mastery learning reduces complications during central venous catheter. *Critical Care Medicine*, *37*, 2697–2701.

Basdogan, C., De, S., Kim, J., Muniyandi, M., Kim, H., & Srinivasan, M. A. (2004). Haptics in minimally invasive surgical simulation. *IEEE Computer Graphics and Applications*, *24*(2), 56–64.

Beaubien, J. M. (2004). The use of simulation for training teamwork skills in health care: How low can you go? *Quality and Safety in Health Care*, *13*(Suppl. 1), i51–i56.

Bell, H., & Waag, W. (1998). Evaluating the effectiveness of flight simulators for training combat skills: A review. *International Journal of Aviation Psychology*, *8*, 223–242.

Berwick, D. M., & Leape, L. L. (1999). Reducing errors in medicine. *Quality in Health Care*, *8*, 145–146.

Blickensderfer, B., Liu, D., & Hernandez, A. (2005). *Simulation-based training: Applying lessons learned in aviation to surface transportation modes*. Daytona Beach, FL: Embry-Riddle Aeronautical University.

Brehmer, B., & Dörner, D. (1993). Experiments with computer-simulated microworlds: Escaping both the narrow straits of the laboratory and the deep blue sea of the field study. *Computers in Human Behavior*, *9*, 171–184.

Bridges, M., & Diamond, D. L. (1999). The financial impact of teaching surgical residents in the operating room. *American Journal of Surgery*, *177*, 28–32.

Caird, J. K. (1996). Persistent issues in the application of virtual environment systems to training. In *Proceedings Third Annual Symposium on Human Interaction with Complex Systems (HICS'96)* (pp. 124–132). Washington, DC: IEEE.

Campbell, S. M., Braspenning, J., Hutchinson, A., & Marshall, M. (2002). Research methods used in developing and applying quality indicators in primary care. *Quality and Safety in Health Care*, *11*, 358–364.

Cant, R. P., & Cooper, S. J. (2010). Simulation-based learning in nurse education: Systematic review. *Journal of Advanced Nursing*, *66*, 3–15.

Carmines, E. G., & Zeller, R. A. (1979). *Reliability and validity assessment*. Thousand Oaks, CA: Sage.

Cepeda, N. J., Pashler, H., Vul, E., Wixted, J. T., & Rohrer, D. (2006). Distributed practice in verbal recall tasks: A review and quantitative synthesis. *Psychological Bulletin*, *132*, 354–380.

Chambers, D. W., Dugoni, A. A., & Paisley, I. (2004). The case against one-shot testing for initial dental licensure. *Journal of the California Dental Association*, *32*, 248–252.

Chipman, S. F., Schraagen, J. M., & Shalin, V. L. (2000). Introduction to cognitive task analysis. In S. F. Chipman, J. M. Schraagen, & V. L. Shalin (Eds.), *Cognitive task analysis* (pp. 3–24). Mahwah, NJ: Lawrence Erlbaum.

Clancy, C. M. (1998). Outcomes research: Measuring the end results of health care. *Science*, *282*, 245–246.

Cohen, E. R., Feinglass, J., Barsuk, J. H., Barnard, C., O'Donnell, A., McGaghie, W. C., & Wayne, D. B. (2010). Cost savings from reduced catheter-related bloodstream infection after simulation-based education for residents in a medical intensive care unit. *Simulation in Healthcare*, *5*, 98–102.

Connolly, T. J., Blackwell, B. B., & Lester, L. (1989). A simulator-based approach to training in aeronautical decision making. *Aviation, Space, and Environmental Medicine*, *60*, 50–52.

Cook, D. A., Hatala, R., & Brydges, R. (2011). Technology-enhanced simulation for health professions education. *Journal of American Medical Association*, *306*, 978–988.

Cook, R. I., & Woods, D. D. (1994). Operating at the sharp end: The complexity of human error. In M. S. Bogner (Ed.), *Human error in medicine* (pp. 255–310), Hillside, NJ: Lawrence Erlbaum.

Cooke, N. J. (1994). Varieties of knowledge elicitation techniques. *International Journal of Human–Computer Studies*, *41*, 801–849.

Cooper, J. B., & Taqueti, V. R. (2004). A brief history of the development of mannequin simulators for clinical education and training. *Quality and Safety Health Care*, *13*(Suppl. 1), i11–i18.

Cosby, K., & Croskerry, P. (2003). Patient safety: A curriculum for teaching patient safety in emergency medicine. *Academic Emergency Medicine*, *10*, 69–78.

Cronbach, L. J., & Meehl, P. E. (1955). Construct validity in psychological tests. *Psychological Bulletin*, *52*, 281–302.

DeAnda, A., & Gaba, D. M. (1990). Unplanned incidents during comprehensive anesthesia simulation. *Anesthesia and Analgesia*, *71*, 77–82.

Delph, S. J., & Zajac, F. R. (1992). Force and moment generating capacity of lower limb muscles before and after tendon lengthening. *Clinical Orthopedics and Related Research*, *284*, 247–259.

DeLucia, P. R., Ott, T. E., & Palmieri, P. A. (2009). Performance in nursing. In F. T. Durso (Ed.), *Reviews of human factors and ergonomics* (Vol. 5, pp. 1–40). Santa Monica, CA: Human Factors and Ergonomics Society.

Devitt, J. H., Kurrek, M. M., Cohen, M. M., & Cleave-Hogg, D. C. (2001). The validity of performance assessments using simulation. *Anesthesiology*, *95*, 36–42.

Dieckmann, P., Gaba, D., & Rall, M. (2007). Deepening the theoretical foundations of patient simulation as a social practice. *Simulation in Healthcare*, *2*, 183–193.

Dillon, G. F. (2004). Simulations in the United States Medical Licensing Examination (USMLE). *Quality and Safety in Health Care*, *13*(Suppl. 1), i41–i45.

Dismukes, R. K., & Smith, G. M. (2001). *Facilitation and debriefing in aviation training and operations*. Aldershot, UK: Ashgate.

Domuracki, K. J., Moule, C. J., Owen, H., Kostandoff, G., & Plummer, J. L. (2009). Learning on a simulator does transfer to clinical practice. *Resuscitation*, *80*, 346–349.

Donchin, Y., Gopher, D., Olin, M., Badihi, Y., Biesky, M. R., Sprung, C. L., & . . . Cotev, S. (1995). A look into the nature and causes of human errors in the intensive care unit. *Critical Care Medicine*, *23*, 294–300.

Dong, Y., Suri, H. S., Cook, D. A., Kashani, K. B., Mullon, J. J., Enders, F. T., & Dunn, W. F. (2010). Simulation-based objective assessment discerns clinical proficiency in central line placement: A construct validation. *Chest*, *137*, 1050–1056.

Draycott, T., Crofts, J., & Ash, J. (2008). Improving neonatal outcome through practical shoulder dystocia training. *Obstetrics & Gynecology*, *112*, 14–20.

Drews, F. A. (2011). Human error. In P. Carayon (Ed.), *Handbook of human factors and ergonomics in healthcare and patient safety* (2nd ed., pp 323–340). Mahwah, NJ: Lawrence Erlbaum.

Drews, F. A. (in press). Adherence engineering. *Ergonomics in Design*.

Drews, F. A., Musters, A., & Samore, M. (2008). Error producing conditions in the intensive care unit. In K. Henriksen, J. B. Battles, M. A. Keyes, & M. L. Grady (Eds.), *Advances in patient safety: From research to implementation. Vol. 5: Performance and tools* (AHRQ Publication No. 050021). Rockville, MD: Agency for Healthcare Research and Quality.

Drews, F. A., Wallace, J., Benuzillo, J., Markewitz, B., & Samore, M. (2012). Protocol adherence in the intensive care unit. *Human Factors and Ergonomics in Manufacturing & Service Industries*, *22*, 21–31.

Drummond, M. F., Sculpher, M. J., Torrance, G. W., O'Brien, B. J., & Stoddart, G. L. (2005). *Methods for the economic evaluation of health care programmes* (3rd ed.). New York, NY: Oxford University Press.

Durso, F. T., & Drews, F. A. (2010). Health care, aviation, and ecosystems: A socio-natural systems perspective. *Current Directions in Psychological Science*, *19*(2), 71–75.

Edelson, D. P., Litzinger, B., Arora, V., Walsh, D., Kim, S., Lauderdale, D. S., & . . . Abella, B. S. (2008). Improving in-hospital cardiac arrest process and outcomes with performance debriefing. *Archives of Internal Medicine*, *168*, 1063–1069.

Ehret, B. D., Gray, W. D., & Kirschenbaum, S. S. (2000). Contending with complexity: Developing and using a scaled world in applied cognitive research. *Human Factors*, *42*, 8–23.

Epstein, R. M. (2002). Defining and assessing professional competence. *Journal of the American Medical Association*, *287*, 226–235.

Ericsson, K. A., Krampe, R. T., & Tesch-romer, C. (1993). The role of deliberate practice in the acquisition of expert performance. *Psychological Review*, *100*, 363–406.

Evans, L. V., & Dodge, K. L. (2010). Simulation and patient safety: Evaluative checklists for central venous catheter insertion. *Quality and Safety in Health Care, 19*(Suppl. 3), i42–i46.

Fann, J. I., Calhoon, J. H., Carpenter, A. J., Merrill, W. H., Brown, J. W., Poston, R. S., & . . . Feins, R. H. (2010). Simulation in coronary artery anastomosis early in cardiothoracic surgical residency training: The boot camp experience. *Journal Thoracic Cardiovasc Surgery*, *139*, 1275–1281.

Fanning, R. M., & Gaba, D. M. (2007). The role of debriefing in simulation-based learning. *Simulation in Healthcare*, *2*, 1–12.

Feldstein, P. J. (2011). *Health care economics*. Clifton Park, NY: Cengage Learning.

Fitch, M. T. (2007). Using high-fidelity emergency simulation with large groups of preclinical medical students in a basic science course. *Medical Teacher*, *29*, 261–263.

Fitts, P. M. (1964). Perceptual-motor skill learning. In A. W. Melton (Ed.), *Categories of human learning* (pp. 243–285). New York, NY: Academic Press.

Flach, J. M., Riccio, G. E., McMillan, G., & Warren, R. (1986). Psychophysical methods for equating performance between alternative motion simulators. *Ergonomics*, *29*, 1423–1438.

Fletcher, G. C., Flin, R. H., McGeorge, P., Glavin, R. J., Maran, N. J., & Patey, R. (2003). Anaesthetists' Non-Technical Skills (ANTS): Evaluation of a behavioural marker system. *British Journal of Anaesthesia*, *90*, 580–588.

Flin, R. H., & Maran, N. J. (2004). Identifying and training non-technical skills for teams in acute medicine. *Quality and Safety in Health Care*, *13*(Suppl. 1), i80–i84.

Fraser, K., Wright, B., Girard, L., Tworek, J., Paget, M., Welikovich, L., & McLaughlin, K. (2011). Simulation training improves diagnostic performance on a real patient with similar clinical findings. *Chest*, *139*, 376–381. doi:10.1378/chest.10-1107

Gaba, D. M. (1989). Human error in anesthetic mishaps. *International Anesthesiology Clinics*, *27*, 137–147.

Gaba, D. M. (1992). Dynamic decision-making in anesthesiology: Cognitive models and training approaches. In D. A. Evans & V. I. Patel (Eds.), *Advanced models of cognition for medical training and practice* (pp. 123–147). Berlin, Germany: Springer-Verlag.

Gaba, D. M. (2004). The future vision of simulation in health care. *Quality and Safety in Health Care*, *13*(Suppl. 1), i2–i10. doi:10.1136/qshc.2004.009878

Gaba, D. M., Howard, S. K., Fish, K. J., Smith, B. E., & Sowb, Y. A. (2001). Simulation-based training in anesthesia crisis resource management (ACRM): A decade of experience. *Simulation & Gaming*, *32*, 175–193.

Gaba, D. M., Howard, S. K., & Small, S. D. (1995). Situation awareness in anesthesiology. *Human Factors*, *37*, 20–31.

Garg, A. X., Adhikari, N. K., McDonald, H., Rosas-Arellano, M. P., Devereaux, P. J., Beyene, J., & . . . Haynes, R. B. (2005). Effects of computerized clinical decision support systems on practitioner performance and patient outcomes: A systematic review. *Journal of the American Medical Association*, *293*, 1223–1238. doi:10.1001/jama.293.10.1223

Gawande, A., (2007). *Complications: A surgeon's notes on an imperfect science*. London, UK: Profile Books.

Gibson, J. J. (1971). The information available in pictures. *Leonardo*, *4*, 27–35.

Gleason, P. L., Kikinis, R., Altobelli, D., Wells, W., Alexander, E., III, Black, P. M., & Jolesz, F. (1995). Video registration virtual reality for nonlinkage stereotactic surgery. *Stereotactic Functional Neurosurgery*, *63*, 139–143.

Good, M. L. (2003). Patient simulation for training basic and advanced clinical skills. *Medical Education*, *37*(Suppl. 1), 14–21.

Gray, W. D. (2002). Simulated task environments: The role of high-fidelity simulations, scaled worlds, synthetic environments, and laboratory tasks in basic and applied cognitive research. *Cognitive Science Quarterly*, *2*, 205–227.

Grenvik, A., & Schaefer, J. (2004). From Resusci-Anne to Sim-Man: The evolution of simulators in medicine. *Critical Care Medicine*, *32*, S56–S57.

Grenvik, A., Schaefer, J. J., DeVita, M. A., & Rogers, P. (2004). New aspects on critical care medicine training. *Current Opinion in Critical Care*, *10*, 233–237.

Grol, R., & Grimshaw, J. (2003). From best evidence to best practice: Effective implementation of change in patients' care. *Lancet*, *362*, 1225–1230.

Haluck, R. A., & Krummel, T. M. (2000). Computers and virtual reality for surgical education in the 21st century. *Archives of Surgery*, *135*, 786–792.

Hamman, W. R. (2004). The complexity of team training: What we have learned from aviation and its applications to medicine. *Quality and Safety in Health Care*, *13*(Suppl. 1), i72–i79.

Hanscom, R. (2008). Medical simulation from an insurer's perspective. *Academic Emergency Medicine*, *15*, 984–987.

Haque, S., & Srinivasan, S. (2006). A meta-analysis of the training effectiveness of virtual reality surgical simulators. *IEEE Transactions on Information Technology in Biomedicine*, *10*, 51–58.

Harden, R. M., Stevenson, M., Downie, W. W., & Wilson, G. M. (1975). Assessment of clinical competence using objective clinical exam. *British Medical Journal*, *1*, 447–451.

Hayden, J. (2010). Use of simulation in nursing education: National survey results. *Journal of Nursing Regulation*, *1*, 52–57.

Hays, E. P. (2010, October). *Advanced wound closure techniques. Paper presented at the American College of Emergency Physicians 2010 Scientific Assembly*, Las Vegas, NV.

Hays, R., Jacobs, J., Prince, C., & Salas, E. (1992). Flight simulator training effectiveness: A meta-analysis. *Military Psychology*, *4*(2), 63–74.

Helmreich, R. L., Wilhelm, J. A., Klinect, J. R., & Merritt, A. C. (2001). Culture, error and crew resource management. In E. Salas, C. A. Bowers, & E. Edens (Eds.), *Applying resource management in organizations: A guide for training professionals* (pp. 305–331). Princeton, NJ: Lawrence Erlbaum.

Hochberg, J. (1986). Representation of motion and space in video and cinematic displays. In K. R. Boff, L. Kaurman, & J. P. Thomas (Eds.), *Handbook of perception and human performance* (Vol. 1, pp. 22–31). New York, NY: Wiley.

Holmboe, E., Rizzolo, M. A., Sachdeva, A. K., Rosenberg, M., & Ziv, A. (2011). Simulation-based assessment and the regulation of healthcare professionals. *Simulation in Healthcare*, *6*, S58–S62.

Hook, J. E.-V. (2004). *Advanced initiatives in medical simulation: Building a national agenda for simulation-based medical education*. Washington, DC: U.S. Army Medical Research and Materiel Command.

Hopkins, C. (1974). How much should you pay for that box? In *Proceedings of the Human Factors and Ergonomics Society 17th Annual Meeting* (pp. 533–541). Santa Monica, CA: Human Factors and Ergonomics Society.

Howard, S., Gaba, D., Fish, K., Yang, G., & Sarnquist, F. (1992). Anesthesia crisis resource management training: Teaching anesthesiologists to handle critical incidents. *Aviation Space and Environmental Medicine*, *63*, 763–770.

Institute for Healthcare Improvement. (n.d.). *The IHI improvement map*. Retrieved from http://www.ihi.org/offerings/Initiatives/Improvemaphospitals/Pages/default.aspx

Issenberg, S. B., McGaghie, W. C., Hart, I. R., Mayer, J. W., Felner, J. M., Petrusa, E. R., & . . . Ewy, G. A. (1999). Simulation technology for health care professional skills training and assessment. *Journal of the American Medical Association*, *282*, 861–866.

Issenberg, S. B., & Scalese, R. J. (2008). Simulation in health care education. *Perspectives in Biology and Medicine*, *51*, 31–46.

Jentsch, F., & Bowers, C. A. (1998). Evidence for the validity of PC-based simulations in studying aircrew coordination. *International Journal of Aviation Psychology*, *8*, 243–260.

Johnson, K., Syroid, N., & Drews, F. (2008). Part task and variable priority training in first-year anesthesia resident education: A combined didactic and simulation-based approach to improve management of adverse airway. *Anesthesiology*, *108*, 831–840.

Johnson, S. J., Guediri, S. M., Kilkenny, C., & Clough, P. J. (2011). Development and validation of a virtual reality simulator: Human factors input to interventional radiology training. *Human Factors*, *53*, 612–625.

Jones, E. R., Hennessey, R. T., & Deutsch, S. (1985). *Human factors aspects of simulation*. Washington, DC: National Research Council.

Kim, H., Rattner, D., & Srinivasan, M. (2003). The role of simulation fidelity in laparoscopic surgical training. In R. E. Ellis & T. M. Peters (Eds.), *MICCAI 2003* (pp. 1–8). Berlin, Germany: Springer-Verlag.

Kim, J., Ritter, F., & Koubek, R. (2013). An integrated theory for improved skill acquisition and retention in the three stages of learning. *Theoretical Issues in Ergonomics Science*, *14*, 22–37.

Klein, M. I., Warm, J. S., Riley, M. A., Matthews, G., Gaitonde, K., & Donovan, J. F. (2008). Perceptual distortions produce multidimensional stress profiles in novice users of an endoscopic surgery simulator. *Human Factors*, *50*, 291–300.

Kneebone, R. (2010). Simulation, safety and surgery. *Quality and Safety in Health Care*, *19*, i47–i52.

Kohn, L. T., Corrigan, J., & Donaldson, M. S. (1999). *To err is human: Building a safer health system* (Vol. 6). Washington, DC: National Academies Press.

Koonce, J. M., & Bramble, W. J. (1998). Personal computer-based flight training devices. *The International Journal of Aviation Psychology*, *8*, 277–292.

Kozlowski, S. W. J., & DeShon, R. P. (2004). A psychological fidelity approach to simulation-based training. In E. Salas, S. G. Schflett, & M. D. Coovert (Eds.), *Scaled worlds: Development, validation, and applications* (pp. 75–99). Burlington, VT: Ashgate.

Lapkin, S., Levett-Jones, T., Bellchambers, H., & Fernandez, R. (2010). Effectiveness of patient simulation manikins in teaching clinical reasoning skills to undergraduate nursing students: A systematic review. *Clinical Simulation in Nursing*, *6*, e207–e222.

Lathan, C. E., Tracey, M. R., Sebrechts, M. M., Clawson, D. M., & Higgins, G. A. (2002). Using virtual environments in training simulators: Measuring transfer. In K. M. Stanney (Ed.), *Handbook of virtual environments: Design, implementation, and applications* (pp. 403–414). Mahwah, NJ: Lawrence Erlbaum.

Leape, L. L. (1994). Error in medicine. *Journal of the American Medical Association*, *272*, 1851–1857.

Lin, M. Y., Hota, B., Khan, Y. M., Woeltje, K. F., Borlawsky, T. B., Doherty, J. A., & . . . Trick, W. E. (2010). Quality of traditional surveillance for public reporting of nosocomial bloodstream infection rates. *Journal of the American Medical Association*, *304*, 2035–2041.

Luck, J., Peabody, J., & Dresselhaus, T. (2000). How well does chart abstraction measure quality? A prospective comparison of standardized patients with the medical record. *American Journal of Medicine*, *108*, 643–649.

Lui, D., Blickensderfer, E. L., Macchiarella, N. D., & Vincenzi, V. A. (2008). Transfer of training. In D. A. Vincenzi, J. A. Wise, M. Mouloua, & P. A. Hancock (Eds.), *Human factors in simulation and training* (pp. 49–60). New York, NY: CRC Press.

Lynagh, M., Burton, R., & Sanson-Fisher, R. (2007). A systematic review of medical skills laboratory training: Where to from here? *Medical Education*, *41*, 879–887.

Makic, M. B. F., VonRueden, K. T., Rauen, C. A., & Chadwick, J. (2011). Evidence-based practice habits: Putting more sacred cows out to pasture. *Critical Care Nurse*, *31*(2), 38–61.

Maran, N. J., & Glavin, R. J. (2003). Low- to high-fidelity simulation: A continuum of medical education? *Medical Education*, *37*(Suppl. 1), 22–28.

Mattoon, J. S. (1994). Designing instructional simulations: Effects of instructional control and type of training task on developing display-interpretation skills. *International Journal of Aviation Psychology*, *4*, 189–209.

McGaghie, W. C., Issenberg, S. B., Cohen, E. R., Barsuk, J. H., & Wayne, D. B. (2011). Does simulation-based medical education with deliberate practice yield better results than traditional clinical education? A meta-analytic comparative review of the evidence. *Academic Medicine*, *86*, 706–711.

McGaghie, W. C., Issenberg, S. B., Petrusa, E. R., & Scalese, R. J. (2010). A critical review of simulation-based medical education research: 2003–2009. *Medical Education*, *44*, 50–63.

McGaghie, W. C., Issenberg, S. B., Cohen, E. R., Barsuk, J. H., & Wayne, D. B. (2011). Does simulation-based medical education with deliberate practice yield better results than traditional clinical education? A meta-analytic comparative review of the evidence. *Academic Medicine*, *86*, 706–711.

Militello, L. G., & Hutton, R. J. B. (1998). Applied cognitive task analysis (ACTA): A practitioner's toolkit for understanding cognitive task demands. *Ergonomics*, *41*, 1618–1641.

Miller, R. B. (1957). *Psychological considerations in the design of training equipment*. Wright-Patterson Air Force Base, OH: Wright Air Development Center.

Moroney, W. F., & Moroney, B. W. (1999). Simulation. In D. J. Garland, J. A. Wise, & V. D. Hopkin (Eds.), *Human factors in aviation systems* (pp. 358–388). Mahwah, NJ: Lawrence Erlbaum.

Morris, C. D., Bransford, J. D., & Franks, J. J. (1977). Levels of processing versus transfer appropriate processing. *Journal of Verbal Learning and Verbal Behavior*, *16*, 519–533.

Morrow, D., North, R., & Wickens, C. D. (2005). Reducing and mitigating human error in medicine. In R. S. Nickerson (Ed.), *Reviews of human factors and ergonomics* (Vol. 1, pp. 254–296). Santa Monica, CA: Human Factors and Ergonomics Society.

Murray, D. J., Boulet, J. R., Avidan, M., Kras, J. F., Henrichs, B., Woodhouse, J., & Evers, A. S. (2007). Performance of residents and anesthesiologists in a simulation-based skill assessment. *Anesthesiology*, *107*, 705–713.

Nehring, W. M. (2008). U.S. boards of nursing and the use of high-fidelity patient simulators in nursing education. *Journal of Professional Nursing*, *24*, 109–117.

Newble, D. I., & Entwistle, N. J. (2009). Learning styles and approaches: Implications for medical education. *Medical Education*, *20*, 162–175.

Okuda, Y., Bryson, E. O., DeMaria, S., Quinones, J., Shen, B., & Levine, A. I. (2009). The utility of simulation in medical education: What is the evidence? *Mount Sinai Journal of Medicine*, *76*, 330–343.

Ortiz, G. A. (1994). Effectiveness of PC-based flight simulation. *International Journal of Aviation Psychology*, *4*, 285–291.

Osgood, C. E. (1949). The similarity paradox in human learning: A resolution. *Psychological Review*, *56*(3), 132–143.

Ost, D., De Rosiers, A., Britt, E. J., Fein, A. M., Lesser, M. L., & Mehta, A. C. (2001). Assessment of a bronchoscopy simulator. *American Journal of Respiratory and Critical Care Medicine*, *164*, 2248–2255.

Page, R. L. (2000). Brief history of flight simulation. In *SimTecT 2000 Proceedings* (pp. 11–17). Lindfield, Australia: Simulation Industry Association of Australia.

Park, J., MacRae, H., Musselman, L. J., Rossos, P., Hamstra, S. J., Wolman, S., & Reznick, R. K. (2007). Randomized controlled trial of virtual reality simulator training: Transfer to live patients. *American Journal of Surgery*, *194*, 205–211.

Pashler, H., & McDaniel, M. (2008). Learning styles concepts and evidence. *Psychological Science in the Public Interest*, *9*, 105–119.

Patel, V. L., Arocha, J. F., & Kaufman, D. R. (2001). A primer on aspects of cognition for medical informatics. *Journal of the American Medical Informatics Association*, *8*, 324–343.

Patrick, J. (1992). *Training: Research and practice*. London, UK: Academic Press.

Pauly-O'Neill, S. (2009). Beyond the five rights: Improving patient safety in pediatric medication administration through simulation. *Clinical Simulation in Nursing*, *5*, e181–e186.

Perkins, G. D. (2007). Simulation in resuscitation training. *Resuscitation*, *73*, 202–211.

Rall, M., & Gaba, D. (2005). Human performance and patient safety. In D. R. Miller (Ed.), *Anesthesia* (6th ed., pp. 3021–3072). Philadelphia, PA: Churchill Livingstone.

Rasmussen, J. (1986). *Information processing and human–machine interaction: An approach to cognitive engineering*. Amsterdam, Netherlands: North-Holland.

Reason, J. (1990). *Human error*. New York, NY: Cambridge University Press.

Reason, J. (1995). Understanding adverse events: Human factors. *Quality in Health Care*, *4*, 80–89.

Rehmann, A. (1995). *A handbook of flight simulation fidelity requirements for human factors research*. Washington, DC: U.S. Department of Transportation.

Riener, C., & Willingham, D. (2010). The myth of learning styles. *Change: The Magazine of Higher Learning*, *42*(5), 32–35.

Rodriguez-Paz, J. M., Kennedy, M., Salas, E., Wu, A. W., Sexton, J. B., Hunt, E. A., & Pronovost, P. J. (2009). Beyond "see one, do one, teach one": Toward a different training paradigm. *Postgraduate Medical Journal*, *85*, 244–249.

Roscoe, S. N. (1991). Simulator qualification: Just as phony as it can be. *International Journal of Aviation Psychology*, *1*, 335–339.

Roscoe, S. N., & Williges, B. (1980). Measurement of the transfer of training. In *Aviation psychology* (pp. 182–193). Ames: Iowa State University Press.

Rosen, J. (1992, June). *From computer-aided design to computer-aided surgery. Paper presented at Medicine Meets Virtual Reality*, San Diego, CA.

Rosen, K. R. (2008). The history of medical simulation. *Journal of Critical Care*, *23*, 157–166.

Rosen, M. A., Salas, E., Wilson, K. A., King, H. B., Salisbury, M., Augenstein, J. S., & . . . Birnbach, D. J. (2008). Measuring team performance in simulation-based training: Adopting best practices for healthcare. *Simulation in Healthcare*, *3*, 33–41.

Ross, A. J., Kodate, N., Anderson, J. E., Thomas, L., & Jaye, P. (2012). Review of simulation studies in anaesthesia journals, 2001–2010: Mapping and content analysis. *British Journal of Anaesthesia*, *109*, 99–109.

Rothschild, J. M., Landrigan, C. P., Cronin, J. W., Kaushal, R., Lockley, S. W., Burdick, E., & . . . Bates, D. W. (2005). The Critical Care Safety Study: The incidence and nature of adverse events and serious medical errors in intensive care. *Critical Care Medicine*, *33*, 1694–1700.

Rudolph, J. W., Simon, R., Raemer, D. B., & Eppich, W. J. (2008). Debriefing as formative assessment: Closing performance gaps in medical education. *Academic Emergency Medicine*, *15*, 1010–1016.

Safar, P., & McMahon, M. (1958). Mouth-to-airway emergency artificial respiration. *Journal of the American Medical Association*, *166*, 1459–1460.

Salas, E., Bowers, C. A., & Rhodenizer, L. (1998). It is not how much you have but how you use it: Toward a rational use of simulation to support aviation training. *International Journal of Aviation Psychology*, *8*, 197–208.

Salas, E., & Burke, C. S. (2002). Simulation for training is effective when . . .. *Quality and Safety in Health Care*, *11*, 119–120.

Salas, E., Tannenbaum, S. I., Kraiger, K., & Smith-Jentsch, K. A. (2012). The science of training and development in organizations: What matters in practice. *Psychological Science in the Public Interest*, *13*, 74–101.

Salas, E., Wilson, K. A., Burke, C. S., & Priest, H. A. (2005). Using simulation-based training to improve patient safety: What does it take? *Journal on Quality and Patient Safety*, *31*, 363–371.

Salas, E., Wilson, K. A., Lazzara, E. H., King, H. B., Augenstein, J., Robinson, D. W., & Bimbach, D. J. (2008). Simulation-based training for patient safety: 10 principles that matter. *Journal of Patient Safety*, *4*, 3–8.

Satava, R. M. (2006). Virtual reality surgical simulator: The first steps. *Clinical Orthopedics and Related Research*, *442*, 2–4.

Satava, R. M., & Jones, S. (2003). Medical applications of virtual reality. In L. J. Hettinger & M. W. Haas (Eds.), *Virtual and adaptive environments* (pp. 325–343). Mahwah, NJ: Lawrence Erlbaum.

Scerbo, M. W., & Dawson, S. (2007). High fidelity, high performance? *Simulation in Healthcare*, *2*, 224–230.

Schlectre, T. M., Bessemer, D. W., & Kolosh, K. P. (1992). Computer-based simulations systems and role-playing: An effective combination for fostering conditional knowledge. *Journal of Computer-Based Instruction*, *19*, 110–114.

Schmidt-Panos, E. A., & Scerbo, M. W. (2008). Simulation advances in medical training laparoscopic skill acquisition. In D. A. Vincenzi, J. A. Wise, M. Mouloua, & P. A. Hancock (Eds.), *Human factors in simulation and training* (pp. 371–386). New York, NY: CRC Press.

Schout, B. M. A., Hendrikx, A. J. M., Scheele, F., Bemelmans, B. L. H., & Scherpbier, A. J. J. A. (2010). Validation and implementation of surgical simulators: A critical review of present, past, and future. *Surgical Endoscopy*, *24*, 536–546.

Scott, R. D. (2009). *The direct medical costs of healthcare-associated infections in U.S. hospitals and the benefits of prevention*. Atlanta, GA: Centers for Disease Control and Prevention.

Sebrechts, M. M., Lathan, C, Clawson, D. M., Miller, M. S., & Trepagnier, C. (2003). Transfer of training in virtual environments: Issues for human performance. In L. J. Hettinger & M. W. Haas (Eds.), *Virtual and adaptive environments* (p. 67–90). Mahwah, NJ: Lawrence Erlbaum.

Seymour, N. E. (2008). VR to OR: A review of the evidence that virtual reality simulation improves operating room performance. *World Journal of Surgery*, *32*, 182–188.

Seymour, N. E., Gallagher, A. G., Roman, S. A., O'Brien, M. K., Bansal, V. K., Andersen, D. K., & Satava, R. M. (2002). Virtual reality training improves operating room performance. *Annals of Surgery*, *236*, 458–464.

Simon, H. A. (1996). *The sciences of the artificial* (3rd ed.). Cambridge, MA: MIT Press.

Singley, M. K., & Anderson, J. R. (1989). *The transfer of cognitive skill*. Boston, MA: Harvard University Press.

Small, S. D., Wuerz, R. C., Simon, R., Shapiro, N., Conn, A., & Setnik, G. (1999). Demonstration of high-fidelity simulation team training for emergency medicine. *Academic Emergency Medicine*, *6*, 312–323.

Smallman, H. S., & St. John, M. S. (2005). Naive realism: Misplaced faith in realistic displays. *Ergonomics in Design*, *13*(3), 6–13.

Smith, N. F., & Smith, D. W. (1989). *Simulators*. New York, NY: Franklin Watts.

Sotomayor, T. M., Salva, A. M., & York, B. W. (2012). Measuring the training effectiveness of combat Lifesaver simulation training systems. In *Proceedings of the Interservice/Industry Training, Simulation & Education Conference (I/ITSEC)*. Arlington, VA: National Training and Simulation Association.

Stappers, P. J. (1997). Intuitive physics in the design of virtual environment interfaces. In G. J. Torenvliet & K. Vicente (Eds.), *Proceedings of the Ninth International Conference on Perception and Action* (pp. 102–106). Toronto, Canada: University of Toronto.

Stappers, P. J. (1998). Experiencing non-Newtonian physics in VR. In M. Bevan (Ed.), *Proceedings of Virtual Reality in Education & Training '97* (pp. 173–179). Loughborough, UK: UR News.

Stappers, P. J., Gaver, G., & Overbeeke, K. (2003). Beyond the limits of real-time realism: Moving from stimulation correspondence to information correspondence. In L. J. Hettinger & M. W. Haas (Eds.), *Virtual and adaptive environments* (pp. 91–110). Mahwah, NJ: Lawrence Erlbaum.

Stappers, P. J., & Waller, P. E. (1993). Using the free fall of objects in visual depth estimation. *Bulletin of the Psychonomic Society*, *31*, 125–127.

Stark, E. A. (1989). Simulation. In R. S. Jensen (Ed.), *Aviation psychology* (pp. 109–153). Brookfield, VT: Gower.

Steadman, R. H., Coates, W. C., Huang, Y. M., Matevosian, R., Larmon, B. R., McCullough, L., & Ariel, D. (2006). Simulation-based training is superior to problem-based learning for the acquisition of critical assessment and management skills. *Critical Care Medicine*, *34*, 151–157.

Stoffregen, T. A., Bardy, B. G., Smart, L. J., & Pagulayan, R. (2003). On the nature and evaluation of fidelity in virtual environments. In L. J. Hettinger & M. W. Haas (Eds.), *Virtual and adaptive environments* (pp. 111–129). Mahwah, NJ: Lawrence Erlbaum.

Strayer, D. L., & Drews, F. A. (2003). Simulator training improves driver efficiency: Transfer from the simulator to the real world. In *Driving Assessment 2003: The Second International Driving Symposium on Human Factors in Driver Assessment, Training and Vehicle Design* (pp. 190–193). Iowa City: University of Iowa.

Strayer, D. L., & Johnston, W. A. (2001). Driven to distraction: Dual-task studies of simulated driving and conversing on a cellular telephone. *Psychological Science*, *12*, 462–466.

Sweller, J. (1994). Cognitive load theory, learning difficulty, and instructional design. *Learning and Instruction*, *4*, 295–312.

Taylor, H. L., Lintern, G., & Koonce, J. M. (1993). Quasi-transfer as a predictor of transfer from simulator to airplane. *Journal of General Psychology*, *120*, 257–276.

Taylor, H. L., Lintern, G., Koonce, J. M., Kaiser, R. H., & Morrison, G. A. (1991, September). *Simulator scene detail and visual augmentation guidance in landing training for beginning pilots. Paper presented at the Aerospace Technology Conference and Exposition*, Long Beach, CA.

Tendick, F., Downes, M., Goktekin, T., Cavusoglu, M. C., Feygin, D., Wu, X., & . . . Way, L. W. (2000). A virtual environment testbed for training laparoscopic surgical skills. *Presence: Teleoperators & Virtual Environments*, *9*, 236–255.

Teteris, E., Fraser, K., Wright, B., & McLaughlin, K. (2012). Does training learners on simulators benefit real patients? *Advances in Medical Education and Practice*, *17*, 137–144.

Timmermans, S., & Berg, M. (2003). *The gold standard: The challenge of evidence-based medicine and standardization in healthcare*. Philadelphia, PA: Temple University Press.

Vozenilek, J., Huff, J. S., Reznek, M., & Gordon, J. A. (2004). See one, do one, teach one: Advanced technology in medical education. *Academic Emergency Medicine*, *11*, 1149–1154.

Wallace, P. (1997). Following the threads of an innovation: The history of standardized patients in medical education. *CADCEUS*, *13*(2), 4–28.

Wass, V., Van der Vleuten, C., Shatzer, J., & Jones, R. (2001). Assessment of clinical competence. *The Lancet*, *357*, 945–949.

Wayne, D. B., Didwania, A., Feinglass, J., Fudala, M. J., Barsuk, J. H., & McGaghie, W. C. (2008). Simulation-based education improves quality of care during cardiac arrest team responses at an academic teaching hospital: A case-control study. *Chest*, *133*, 56–61.

Wayne, D. B., Siddall, V. J., Butter, J., Fudala, M. J., Wade, L. D., Feinglass, J., & McGaghie, W. C. (2006). A longitudinal study of internal medicine residents' retention of advanced cardiac life support skills. *Academic Medicine*, *81*(10), S9–S12.

Weaver, S. J., Salas, E., Lyons, R., Lazzara, E. H., Rosen, M. A., Diazgranados, D., & . . . King, H. (2010). Simulation-based team training at the sharp end: A qualitative study of simulation-based team training design, implementation, and evaluation in healthcare. *Journal of Emergencies: Trauma and Shock*, *3*, 369–377. doi:10.4103/0974-2700.70754

Wickens, C. D., Gordon, S. E., & Liu, Y. (1998). *An introduction to human factors engineering*. New York, NY: Addison Wesley Educational.

Witmer, B., & Singer, M. (1998). Measuring presence in virtual environments: A presence questionnaire. *Presence*, *7*, 225–240.

Woolliscroft, J. O., Calhoun, J. G., Tenhaken, J. D., & Judge, R. D. (1987). Harvey: The impact of a cardiovascular teaching simulator on student skill acquisition. *Medical Teacher*, *9*, 53–57.

Xiao, Y., Parker, S. H., & Manser, T. (in press). Teamwork and collaboration. In D. G. Morrow (Ed.), *Reviews of human factors and ergonomics* (Vol. 8). Santa Monica, CA: Human Factors and Ergonomics Society.

Yee, B., Naik, V. N., Joo, H. S., Savoldelli, G. L., Chung, D. Y., Houston, P. L., & . . . Hamstra, S. J. (2005). Nontechnical skills in anesthesia crisis management with repeated exposure to simulation-based education. *Anesthesiology*, *103*, 241–248.

Zendejas, B., Wang, A. T., Brydges, R., Hamstra, S. J., & Cook, D. A. (2013). Cost: The missing outcome in simulation-based medical education research: A systematic review. *Surgery*, *153*, 160–176.

Zhang, Q., & Li, B. (2010). Towards computational understanding of skill levels in simulation-based surgical training via automatic video analysis. In G. Bebis et al. (Eds.), *ISVC 2010, Part III, LNCS 6455* (pp. 249–260). Berlin, Germany: Springer-Verlag.

Ziv, A., Wolpe, P. R., Small, S. D., & Glick, S. (2006). Simulation-based medical education: An ethical imperative. *Simulation in Healthcare*, *1*, 252–256.

ABOUT THE AUTHORS

Frank A. Drews is an associate professor of psychology at the University of Utah and the director of the Center for Human Factors in Patient Safety at the George Wahlen VA Medical Center in Salt Lake City, Utah. He received his PhD in psychology from the Technical University of Berlin, Germany, in 1999.

Jonathan Z. Bakdash is a research psychologist at the U.S. Army Research Laboratory. He received his PhD in psychology from the University of Virginia in 2010.

1 (PDF)	DEFENSE TECHNICAL INFORMATION CTR DTIC OCA
2 (PDF)	DIRECTOR US ARMY RESEARCH LAB RDRL CIO LL IMAL HRA MAIL & RECORDS MGMT
1 (PDF)	GOVT PRINTG OFC A MALHOTRA
1 (PDF)	ARMY RSCH LABORATORY – HRED RDRL HRM D T DAVIS BLDG 5400 RM C242 REDSTONE ARSENAL AL 35898-7290
1 (PDF)	ARMY RSCH LABORATORY – HRED RDRL HRS EA DR V J RICE BLDG 4011 RM 217 1750 GREELEY RD FORT SAM HOUSTON TX 78234-5002
1 (PDF)	ARMY RSCH LABORATORY – HRED RDRL HRM DG J RUBINSTEIN BLDG 333 PICATINNY ARSENAL NJ 07806-5000
1 (PDF)	ARMY RSCH LABORATORY – HRED ARMC FIELD ELEMENT RDRL HRM CH C BURNS THIRD AVE BLDG 1467B RM 336 FORT KNOX KY 40121
1 (PDF)	ARMY RSCH LABORATORY – HRED AWC FIELD ELEMENT RDRL HRM DJ D DURBIN BLDG 4506 (DCD) RM 107 FORT RUCKER AL 36362-5000
1 (PDF)	ARMY RSCH LABORATORY – HRED RDRL HRM CK J REINHART 10125 KINGMAN RD BLDG 317 FORT BELVOIR VA 22060-5828
1 (PDF)	ARMY RSCH LABORATORY – HRED RDRL HRM AY M BARNES 2520 HEALY AVE STE 1172 BLDG 51005 FORT HUACHUCA AZ 85613-7069

1 (PDF)	ARMY RSCH LABORATORY – HRED RDRL HRM AP D UNGVARSKY POPE HALL BLDG 470 BCBL 806 HARRISON DR FORT LEAVENWORTH KS 66027-2302
1 (PDF)	ARMY RSCH LABORATORY – HRED RDRL HRM AT J CHEN 12423 RESEARCH PKWY ORLANDO FL 32826-3276
1 (PDF)	ARMY RSCH LAB – HRED HUMAN SYSTEMS INTEGRATION ENGR TACOM FIELD ELEMENT RDRL HRM CU P MUNYA 6501 E 11 MILE RD MS 284 BLDG 200A WARREN MI 48397-5000
1 (PDF)	ARMY RSCH LABORATORY – HRED FIRES CTR OF EXCELLENCE FIELD ELEMENT RDRL HRM AF C HERNANDEZ 3040 NW AUSTIN RD RM 221 FORT SILL OK 73503-9043
1 (PDF)	ARMY RSCH LABORATORY – HRED RDRL HRM AV W CULBERTSON 91012 STATION AVE FORT HOOD TX 76544-5073
1 (PDF)	ARMY RSCH LABORATORY – HRED RDRL HRM DE A MARES 1733 PLEASONTON RD BOX 3 FORT BLISS TX 79916-6816
8 (PDF)	ARMY RSCH LABORATORY – HRED SIMULATION & TRAINING TECHNOLOGY CENTER RDRL HRT COL G LAASE RDRL HRT I MARTINEZ RDRL HRT T R SOTTILARE RDRL HRT B N FINKELSTEIN RDRL HRT G A RODRIGUEZ RDRL HRT I J HART RDRL HRT M C METEVIER RDRL HRT S B PETTIT 12423 RESEARCH PARKWAY ORLANDO FL 32826

1   ARMY RSCH LABORATORY – HRED
(PDF)   HQ USASOC
         RDRL HRM CN   R SPENCER
         BLDG E2929 DESERT STORM DRIVE
         FORT BRAGG NC 28310

1   ARMY G1
(PDF)   DAPE MR   B KNAPP
         300 ARMY PENTAGON RM 2C489
         WASHINGTON DC 20310-0300

         ABERDEEN PROVING GROUND

12   DIR USARL
(PDF)   RDRL HR
            L ALLENDER
            P FRANASZCZUK
            K MCDOWELL
         RDRL HRM
            P SAVAGE-KNEPSHIELD
         RDRL HRM AL
            C PAULILLO
         RDRL HRM B
            J GRYNOVICKI
         RDRL HRM C
            L GARRETT
         RDRL HRS
            J LOCKETT
         RDRL HRS B
            M LAFIANDRA
         RDRL HRS D
            A SCHARINE
         RDRL HRS E
            J BAKDASH
            D HEADLEY

1   UNIV UTAH
(PDF)   F DREWS